



(19) 中華民國智慧財產局

(12) 發明說明書公開本

(11) 公開編號：TW 201730875 A

(43) 公開日：中華民國 106 (2017) 年 09 月 01 日

(21) 申請案號：106115507

(22) 申請日：中華民國 103 (2014) 年 01 月 16 日

(51) Int. Cl. : G10L19/00 (2013.01)

(30) 優先權：2013/01/21 美國 61/754,882

2013/05/16 美國 61/824,010

(71) 申請人：杜比實驗室特許公司(美國) DOLBY LABORATORIES LICENSING CORPORATION
(US)

美國

(72) 發明人：葛蘭特 麥可 GRANT, MICHAEL (AU)；諾克羅斯 史考特 NORCROSS, SCOTT
GREGORY (CA)；萊德米勒 傑佛瑞 RIEDMILLER, JEFFREY C. (US)；沃德 麥
可 WARD, MICHAEL (GB)

(74) 代理人：林志剛

申請實體審查：有 申請專利範圍項數：7 項 圖式數：11 共 92 頁

(54) 名稱

使用位在保留資料空間中的元資料容器解碼已編碼音頻位元流

DECODING OF ENCODED AUDIO BITSTREAM WITH METADATA CONTAINER LOCATED IN
RESERVED DATA SPACE

(57) 摘要

本發明揭露用來產生一已編碼音頻位元流的設備與方法，其包括藉由包括節目響度元資料與音頻資料於該位元流中，以及可選地亦包括節目邊界元資料於該位元流的至少一個片段(例如，訊框)中。其他態樣係為用來解碼此類位元流的設備與方法，例如包括藉由施行由該位元流所指示出之一音頻節目的該音頻資料的自適應響度處理，或者此類音頻節目之元資料及/或音頻資料的鑑定及/或驗證。另一態樣係為一音頻處理單元(例如，一編碼器、解碼器或後處理器)，其被組態(例如，程式化)以施行該方法之任何實施例或者其包括儲存根據該方法之任何實施例所產生之一音頻位元流之至少一個訊框的一緩衝記憶體。

Apparatus and methods for generating an encoded audio bitstream, including by including program loudness metadata and audio data in the bitstream, and optionally also program boundary metadata in at least one segment (e.g., frame) of the bitstream. Other aspects are apparatus and methods for decoding such a bitstream, e.g., including by performing adaptive loudness processing of the audio data of an audio program indicated by the bitstream, or authentication and/or validation of metadata and/or audio data of such an audio program. Another aspect is an audio processing unit (e.g., an encoder, decoder, or post-processor) configured (e.g., programmed) to perform any embodiment of the method or which includes a buffer memory which stores at least one frame of an audio bitstream generated in accordance with any embodiment of the method.

指定代表圖：

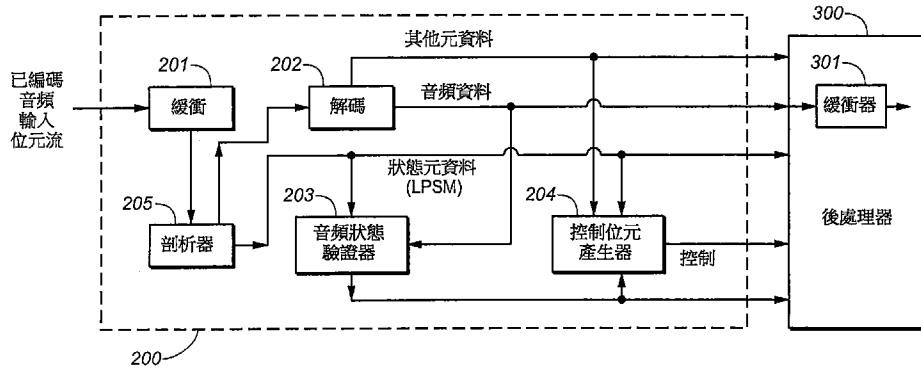


圖 3

符號簡單說明：

200 . . . 解碼器

201 . . . 訊框緩衝器

202 . . . 音頻解碼器

203 . . . 音頻狀態驗證台(驗證器)

204 . . . 控制位元產生台

205 . . . 剖析器

300 . . . 後處理器

301 . . . 訊框緩衝器

發明摘要

※申請案號： 106115507 (由106112038分割)

※申請日： 103/01/16

※IPC 分類： *G10L 19/00* (2013.01)

【發明名稱】 (中文/英文)

使用位在保留資料空間中的元資料容器解碼已編碼音頻位元流

Decoding of encoded audio bistream with metadata container located in reserved data space

【中文】

本發明揭露用來產生一已編碼音頻位元流的設備與方法，其包括藉由包括節目響度元資料與音頻資料於該位元流中，以及可選地亦包括節目邊界元資料於該位元流的至少一個片段（例如，訊框）中。其他態樣係為用來解碼此類位元流的設備與方法，例如包括藉由施行由該位元流所指示出之一音頻節目的該音頻資料的自適應響度處理，或者此類音頻節目之元資料及/或音頻資料的鑑定及/或驗證。另一態樣係為一音頻處理單元（例如，一編碼器、解碼器或後處理器），其被組態（例如，程式化）以施行該方法之任何實施例或者其包括儲存根據該方法之任何實施例所產生之一音頻位元流之至少一個訊框的一緩衝記憶體。

【 英文 】

Apparatus and methods for generating an encoded audio bitstream, including by including program loudness metadata and audio data in the bitstream, and optionally also program boundary metadata in at least one segment (e.g., frame) of the bitstream. Other aspects are apparatus and methods for decoding such a bitstream, e.g., including by performing adaptive loudness processing of the audio data of an audio program indicated by the bitstream, or authentication and/or validation of metadata and/or audio data of such an audio program. Another aspect is an audio processing unit (e.g., an encoder, decoder, or post-processor) configured (e.g., programmed) to perform any embodiment of the method or which includes a buffer memory which stores at least one frame of an audio bitstream generated in accordance with any embodiment of the method.

【代表圖】

【本案指定代表圖】：第(3)圖。

【本代表圖之符號簡單說明】：

200：解碼器

201：訊框緩衝器

202：音頻解碼器

203：音頻狀態驗證台（驗證器）

204：控制位元產生台

205：剖析器

300：後處理器

301：訊框緩衝器

【本案若有化學式時，請揭示最能顯示發明特徵的化學式】：無

發明專利說明書

(本說明書格式、順序，請勿任意更動)

【發明名稱】(中文/英文)

使用位在保留資料空間中的元資料容器解碼已編碼音頻位元流

Decoding of encoded audio bistream with metadata container located in reserved data space

【技術領域】

本發明關於音頻訊號處理，且更特別地關於音頻資料位元流之編碼與解碼，其具有指示出音頻內容之響度處理狀態以及由該等位元流所指示之音頻節目邊界的位置的元資料。本發明的一些實施例將呈已知為 AC-3、增強 AC-3、或 E-AC-3 或杜比 E 之該等格式之其中一個的音頻資料產生或解碼。

【先前技術】

杜比、杜比數位、杜比數位 Plus、與杜比 E 係為杜比實驗室特許公司的商標。杜比實驗室提供個別已知為杜比數位與杜比數位 Plus 之 AC-3 與 E-AC-3 的專屬實施程序。

音頻資料處理單元基本上以盲目的方式操作，並且沒有注意到在接收資料以前所發生之音頻資料的處理歷史。這或許可在單一個體進行全部的音頻資料處理與編碼，以

用於各種目標媒體顯像裝置，同時目標媒體顯像裝置進行已編碼音頻資料的全部解碼與顯像的處理架構中運行。不過，此盲目處理在複數個音頻處理單元被分散在不同網路或者串聯地放置（亦即，鏈）以及被期待最佳施行它們各別類型的音頻處理的情況下沒有（或者一點都沒有）運行良好。例如，一些音頻資料可被編碼以用於高性能媒體系統，並且必須被轉換成適合用於沿著一媒體處理鏈之行動裝置的一縮減形式。據此，音頻處理單元可非必要地施行一種處理於已經被施行的音頻資料上。例如，音量分級單元可施行處理於輸入音頻段落（clip）上，不管相同或相似的音量分級是否已經被事先施行於輸入音頻段落上。因此，音量分級單元甚至在不必要時仍施行分級。此非必要處理亦會在顯現音頻資料的內容的同時，導致具體特徵的退化與/或移除。

典型的音頻資料流包括指示出該音頻內容之至少一個特徵的音頻內容（例如，音頻內容的一或多個頻道）與元資料兩者。例如，在 AC-3 位元流中，有數個音頻元資料參數，其係特別打算使用來改變被傳送到聆聽環境之節目的聲音。該等元資料參數的其中一個為對話歸一化（DIALNORM）參數，其係打算指出發生音頻節目之對話的平均位準，並且被使用來決定音頻播放訊號位準。

在包含一序列不同音頻節目片段之位元流的播放期間內（各個均具有不同的 DIALNORM 參數），AC-3 解碼器使用各片段的 DIALNORM 參數來施行一種響度處理，其

中它修改播放位準或響度，以致於該序列片段之對話的接收響度是在一致的位準上。在一序列已編碼音頻項目中的各已編碼音頻片段（項目）將（通常）具有不同的 DIALNORM 參數，且該解碼器將按比例縮放各項目的位準，使得各項目之對話的播放位準或響度相同或非常類似，雖然這需要在播放期間內將不同數量的增益施加到該等項目的不同者。

DIALNORM 基本上是由使用者所設定，而且不會自動產生，雖然若使用者沒有設定任何值，則會有一預設的 DIALNORM 值。例如，一內容產生器可用 AC-3 編碼器外部的裝置來進行響度測量，並且隨後將該結果（指示出音頻節目之口語對話的響度）傳送到該編碼器，以設定 DIALNORM 值。因此可依靠該內容產生器，以正確地設定 DIALNORM 參數。

為何在 AC-3 位元流中的 DIALNORM 參數會不正確，有數個不同因素。首先，假如 DIALNORM 值不由內容產生器設定，各 AC-3 編碼器會具有在位元流產生期間內被使用的預設 DIALNORM 值。此預設值會與該音頻的真實對話響度位準實質不同。第二，即使一內容產生器測量響度並且據此設定 DIALNORM 值，一響度測量演算法或計則會被使用，而其並不符合建議的 AC-3 響度測量方法，因而造成不正確的 DIALNORM 值。第三，即使 AC-3 位元流已經用由內容產生器所正確測量與設定的 DIALNORM 值來產生，它可在該位元流的傳輸及/或儲存

期間內被改變為一不正確的值。例如，在電視廣播應用中，解碼、修改並且隨後使用不正確的 DIALNORM 元資料資訊重新解碼 AC-3 位元流並不罕見。因此，AC-3 位元流中所包括的 DIALNORM 值可能係不正確或不準確的，並且因此在聆聽經驗的品質上具有負面影響。

再者，DIALNORM 參數並不指示對應音頻資料的響度處理狀態（例如，哪（些）類型的響度處理已經在音頻資料上被施行）。直到本發明，音頻位元流不曾包括元資料，其指示響度處理狀態（例如，施加到該音頻位元流之音頻內容的響度處理類型），或者該位元流之音頻內容的響度處理狀態與響度，其係呈本發明所說明類型的格式。以特別有效的方式，呈此類格式的響度處理狀態元資料，其對促進音頻位元流之自適應響度處理及/或該音頻內容之響度處理狀態與響度之效度的檢驗有用。

雖然本發明不限於使用 AC-3 位元流、E-AC-3 位元流或杜比 E 位元流，為了方便起見，它將在它產生、解碼或者另外處理此類位元流的實施例中被說明，該位元流包括響度處理狀態元資料。

AC-3 已編碼位元流包含元資料以及一至六個頻道的音頻內容。該音頻內容為已經使用感知音頻編碼來壓縮的音頻資料。該元資料包括數個音頻元資料參數，其打算使用來改變被傳送到聆聽環境之節目的聲音。

AC-3（亦已知為杜比訊號）編碼的細節為眾所皆知，並且陳述許多公布的參考，包括以下：

ATSC 標準 A52/A：數位音頻壓縮標準（AC-3），修訂版 A，高級電視系統委員會，2001 年 8 月 20 日；以及
美國專利 5,583,962；5,632,005；5,633,981；
5,727,119 與 6,021,386，全部在此以引用的方式併入本文。

杜比數位 Plus（E-AC-3）編碼的細節係被陳述在“對杜比數位 Plus 的介紹，杜比數位編碼系統的增加” AES 會議論文 6196，第 117 次 AES 大會，2004 年 10 月 28 日。

杜比 E 編碼的細節係被陳述在“在音頻分佈系統中的有效位元配置、量化與編碼”，AES Preprint 5068，第 107 子 AES 會議，1999 年 8 月，以及“使用視頻來最佳化的專業音頻編碼器”，AES Preprint 5033，第 107 子 AES 會議，1999 年 8 月。

AC-3 已編碼音頻位元流的各訊框包含 1536 個數位音頻樣本的音頻內容與元資料。就 48kHz 的取樣速率而言，這代表 32 毫秒的數位音頻或者每秒音頻 31.25 訊框的速率。

E-AC-3 已編碼音頻位元流的各訊框包含 256,512,768 的音頻內容與元資料或者 1536 個數位音頻樣本，其係取決於該訊框是否各自包含一個、兩個、三個或六個方塊的音頻資料。就 48kHz 的取樣速率而言，這各別代表 5.333、10.667、16 或 32 毫秒的數位音頻，或者各別代表每秒音頻 189.9、93.75、62.5 或 31.25 訊框。

如圖 4 所指示，各 AC-3 訊框會被分成段（片段），其包括：同步資訊（SI）段，其包含（如圖 5 所示）同步字（SW）與兩個錯誤校正字（CRC1）的第一個；包含大部分元資料的位元流資訊（BSI）段；六個音頻方塊（AB0 至 AB5），其包含資料壓縮音頻內容（並且亦包括元資料）；無用位元片段（W），其包含在該音頻內容被壓縮以後殘餘的任何不用位元；一輔助（AUX）資訊段，其包含更多元資料；以及兩個錯誤校正字的第二個（CRC2）。無用位元片段（W）亦稱為“跨越欄位”。

如圖 7 所指示，各 E-AC-3 訊框被分成段（片段），包括：同步資訊（SI）段，其包含（如圖 5 所示）同步字（SW）；包含大部分元資料的位元流資訊（BSI）段；在一個與六個音頻方塊（AB0 至 AB5）之間，其包含資料壓縮音頻內容（並且亦包括元資料）；無用位元片段（W），其包含在該音頻內容被壓縮以後殘餘的任何不用位元（雖然只有一個無用位元片段被顯示，但是不同無用位元片段基本上將跟隨各音頻方塊）；一輔助（AUX）資訊段，其包含更多元資料；以及一錯誤校正字（CRC）。無用位元片段（W）亦稱為“跨越欄位”。

在 AC-3（或 E-AC-3）位元流中，有數個音頻元資料參數，其係特別打算使用來改變被傳送到聆聽環境之節目的聲音。該等元資料參數的其中一個為 DIALNORM 參數，其被包括在 BSI 片段中。

如圖 6 所示，AC-3 訊框的 BSI 片段包括五位元參數

("DIALNORM")，其指示該節目的 DIALNORM 值。若 AC-3 訊框的音頻編碼模組 ("acmod") 是 "0"，指示在相同 AC-3 訊框中所攜帶之第二音頻節目的 DIALNORM 值的五位元參數 ("DIALNORM2") 係被包括，其指示雙-單或 "1+1" 頻道組態在使用中。

BSI 片段亦包括旗標 ("addbsie")，其指示接著 "addbsie" 位元之後的額外位元流資訊的存在 (或缺席)；一參數 ("addbsil")，其指示接著 "addbsil" 值之後的任何額外位元流資訊的長度；以及多達 64 位元的額外位元流資訊 ("addbsi")，其接著 "addbsil" 值之後。

BSI 片段包括沒有特別顯示於圖 6 的其他元資料值。

【發明內容】

在一種類的實施例中，本發明為一音頻處理單元，其包括緩衝器記憶體、音頻解碼器與剖析器。該緩衝器記憶體儲存一已編碼音頻位元流的至少一個訊框。該已編碼音頻位元流包括音頻資料與元資料容器。該元資料容器包括一標頭、一或多個元資料酬載、與保護資料。該標頭包括識別該容器之開頭的一同步字。該一或多個元資料酬載說明與該音頻資料有關的一音頻節目。該保護資料放置在該一或多個元資料酬載以後。該保護資料亦能夠被使用來驗證該元資料容器以及在該元資料容器內之該一或多個酬載的完整性。該音頻解碼器係耦合到該緩衝記憶體，以及能

夠解碼該音頻資料。該剖析器係耦合到或整合該音頻解碼器，以及能夠剖析該元資料容器。

在典型的實施例中，該方法包括接收一已編碼音頻位元流，在此，該已編碼音頻位元流被分段為一或多個訊框。該音頻資料係從該已編碼音頻位元流擷取，連同元資料的一容器。該元資料的容器包括一標頭，其後接著一或多個元資料酬載，其後接著保護資料。最後，該容器與該一或多個元資料酬載的完整性係經由使用該保護資料來檢驗。該一或多個元資料酬載包括一節目響度酬載，該節目響度酬載包含指示與該音頻資料有關之一音頻節目之該測量響度的資料。

節目響度元資料的酬載，稱為響度處理狀態元資料（“LPSM”），根據本發明的典型實施例被嵌入於音頻位元流中，其係可被鑑定與驗證，例如，以致使響度管制實體檢驗一特定節目的響度是否已經在一特定的範圍內，且該相應音頻資料本身則不會被修改（因而確保遵守可應用的規則）。在包含響度處理狀態元資料之資料方塊中所包括的響度值可被讀出，以檢驗此，而不是再度計算響度。回應 LPSM，管制媒介決定對應的音頻內容遵守（如 LPSM 所指示）響度法定及/或管制規格，（例如，該等規則係在商業廣告音量調降法案（Commercial Advertisement Loudness Mitigation Act）之下公布，亦視為“CALM”法案，其係不需要計算該音頻內容的響度。

為了遵守一些響度法定及/或管制規格所必要的響度

測量（例如，該等規則係在 CALM 法案之下公布），其係以整合節目響度為基礎。整合節目響度要求對話位準或全混合位準的響度測量在整個音頻節目上進行。因此，為了進行節目響度測量（例如，在廣播鏈中的各階段上），以檢驗對基本法定規格的遵守，以對哪種音頻資料（與元資料）決定整個音頻節目的知悉來進行決定是重要的，而且這基本上需要對該節目開始與結束之位置的知悉（例如，在指示一序列音頻節目之位元流的處理期間內）。

根據本發明的典型實施例，一已編碼音頻位元流指示至少一個音頻節目（例如，一序列的音頻節目），而且被包括在該位元流中的節目邊界元資料與 LPSM 則致使節目響度測量在節目結束時重新設定，並因而提供一種測量整體節目響度的自動化方式。本發明的典型實施例包括在有效方式中之已編碼音頻位元流中的節目邊界元資料，其允許在該位元流所指示之連續音頻節目之間之至少一個邊界的準確與穩定決定。典型的實施例在它們允許準確節目邊界決定的意義中允許一節目邊界準確且穩定的決定，甚至假使其中指示不同節目的位元流以截斷該等接合位元流其中一個或兩個的方式被接合在一起（以產生本發明位元流）（並且因此放棄已經被包括在至少一個事先接合位元流中的節目邊界元資料）。

在典型的實施例中，在本發明位元流訊框中的節目邊界元資料為指示訊框數的節目邊界旗標。基本上，該旗標指示在目前訊框（包括該旗標的該訊框）與節目邊界（該

目前音頻節目的開始或結束)之間的訊框數目。在一些較佳實施例中，節目邊界旗標以對稱、有效的方式在指示單一節目之各位元流片段開始與結束時被插入(亦即，在該片段開始以後之發生於一些預定數目訊框內的訊框中，以及在該片段結束以前之發生於一些預定數目訊框內的訊框中)，以致於當兩個此位元流片段被連接時(以便指示一兩個節目的序列)，該節目邊界元資料可(例如，對稱地)存在於該兩節目之間之邊界的兩側上。

為了限制由於將節目邊界元資料加入已編碼音頻位元流(其指示一個音頻節目或者一序列音頻節目)中所造成的資料速率增加，在典型的實施例中，僅於該位元流之該等訊框的一子集中節目邊界旗標。基本上，該邊界旗標插入速率為各位元流訊框(其中插入一旗標)距最近該各訊框之該節目邊界之增加間距的非增加函數，在此“邊界旗標插入速率”指示包括節目邊界旗標之訊框數目(指示一節目)對不包括節目邊界旗標之訊框數目(指示該節目)的平均比率，在此，該平均值為已編碼音頻位元流之若干(例如，相當小數目)連續訊框上的移動平均值。在一種類的實施例中，該邊界旗標插入速率為距該最近節目邊界之(各旗標插入位置之)增加距離的對數減少函數，而且就包括其中一個旗標的各包含旗標訊框而言，在該包含旗標訊框中之旗標的尺寸等於或大於在位置比該包含旗標訊框還更靠近最近節目邊界之訊框中各旗標的尺寸(亦即，在各包含旗標訊框中之節目邊界旗標的尺寸，其為距該最

近節目邊界之該包含旗標訊框之增加間隔的非減少函數)。

本發明的另一種態樣為被組態以施行本發明方法之任何實施例的音頻處理單元 (APU)。在另一種類的實施例中，本發明為一 APU，其包括將由本發明方法之任何實施例所產生之已編碼音頻位元流之至少一個訊框儲存 (例如，以非過渡的方式) 的緩衝器記憶體 (緩衝器)。APU 的實例包括但不限於編碼器 (例如，轉碼器)、解碼器、編解碼器、預處理系統 (預處理器)、後處理系統 (後處理器)、音頻位元流處理系統與此類元件之組合。

在另一種類的實施例中，本發明為被組態以產生包含音頻資料片段與元資料片段之已編碼音頻位元流的音頻處理單元 (APU)，在此該音頻資料片段指示音頻資料，且至少一些元資料片段的各個則包括響度處理狀態元資料 (LPSM) 以及可選地亦包括節目邊界元資料。基本上，在該位元流之訊框中的至少一個此元資料片段包括 LPSM 的至少一個片段，其指示第一種類的響度處理是否已經被施行於該訊框的音頻資料上 (亦即，在該訊框之至少一個音頻資料片段中的音頻資料)，以及 LPSM 的至少一個其他片段，其指示該訊框之至少一些音頻資料的響度 (例如，指示對話之該訊框之至少一些音頻資料的對話響度)。在此種類的一項實施例中，APU 為被組態以編碼輸入音頻以產生已編碼音頻的編碼器，而且該音頻資料片段包括已編碼音頻。在此種類的典型實施例中，該等元資料

片段的各個具有在本文中所說明的較佳格式。

在一些實施例中，包括 LPSM（例如，LPSM 與節目邊界元資料）之已編碼位元流（在一些實施例中的 AC-3 位元流或 E-AC-3 位元流）的各元資料片段，係被包括在該位元流之一訊框之跨越欄位片段的無用位元中（例如，顯示於圖 4 或圖 7 之該種類的無用位元片段 W）。在其他實施例中，包括 LPSM（例如，LPSM 與節目邊界元資料）之已編碼位元流（在一些實施例中的 AC-3 位元流或 E-AC-3 位元流）的各元資料片段，其係當作額外位元流資訊地被包括在該位元流之一訊框之位元流資訊（“BSI”）片段的“addbsi”欄位中或者在該位元流之一訊框結束時的輔助資料欄位中（例如，顯示於圖 4 或圖 7 之該種類的 AUX 片段）。包括 LPSM 的各元資料片段具有參考下文表 1 與 2 之在本文中所具體指定的格式（亦即，它包括在表 1 中所具體指定的核心元件或其變化），其後接著酬載 ID（將該元資料識別為 LPSM）與酬載尺寸值，其後接著酬載（具有如表 2 中所指示格式或者如本文中所說明之表 2 之變化所指示之格式的 LPSM 資料）。在一些實施例中，一訊框包括一或兩個元資料片段，其中各個包括 LPSM，且假如該訊框包括兩個元資料片段，一個存在於該訊框的 addbsi 欄位，且另一個在該訊框的 AUX 欄位中。

在一種類的實施例中，本發明為包括將音頻資料編碼以產生 AC-3 或 E-AC-3 已編碼音頻位元流之步驟的方

法，其包括藉由包括在元資料片段（該位元流之至少一個訊框的）LPSM 與節目邊界元資料與可選地亦包括該訊框所屬之音頻節目用的其他元資料。在一些實施例中，各此元資料片段被包括在該訊框的額外位元流（adbsi）欄位，或者該訊框的輔助資料（auxdata）欄位中。在其他實施例中，各此類的元資料片段被包括在該訊框的無用位元片段中。在一些實施例中，包含 LPSM 與節目邊界元資料的各元資料片段包含一核心標頭（以及可選地亦包括額外核心元件），以及在該核心標頭（或者該核心標頭與其他核心元件）以後，LPSM 酬載（或容器）片段具有以下格式：

一標頭，基本上包括至少一個識別值（例如，LPSM 格式版本、長度、時期、數目與子串流相關值，如在本文所陳述的表 2 中指出），以及

在該標頭以後，LPSM 與節目邊界元資料。該節目邊界元資料包括一節目邊界訊框數目，以及一代碼值（例如，一“偏移_存在”值），其指示該訊框是否僅包括一節目邊界訊框數或者一節目邊界訊框數與一偏移值兩者，以及（在一些情形中）一偏移值。LPSM 包括：

至少一個對話指示值，其指示對應的音頻資料是否指示對話或不指示對話（例如，對應音頻資料的哪些頻道指示對話）。該對話指示值可指示對話是否存在於該對應音頻資料之頻道的任一組合或全部中；

至少一個響度規則遵守值，指示對應的音頻資料是否

遵守一指定組響度規則；

至少一個響度處理值，指示至少一種響度處理，其已經在該對應的音頻資料上施行；以及

至少一個響度值，指示該對應音頻資料的至少一個響度（例如，峰值或平均響度）特徵。

在其他實施例中，已編碼位元流為非 AC-3 位元流或 E-AC-3 位元流之一位元流，且包括 LPSM（以及可選地亦包括節目邊界元資料）之元資料片段的各個被包括在為了儲存額外資料而保留之位元流的片段（或者欄位或槽）中。包括 LPSM 的各元資料片段具有與參考下文表 1 與 2 而在本文中所具體指定者相似或相同的格式（亦即，它包括與那些在表 1 所具體指定者類似或相同的核心元件，其後接著酬載 ID（將該元資料識別為 LPSM）與酬載尺寸值，其後接著酬載（具有與在表 2 所指示格式類似或相同的格式或者本文中所說明之表 2 的變化的 LPSM 資料））。

在一些實施例中，已編碼位元流包含一序列的訊框，該訊框的各個均包括位元流資訊（“BSI”）片段，其包括“adbsi”欄位（有時稱為片段或槽）及輔助資料欄位或槽（例如，已編碼位元流為 AC-3 位元流或 E-AC-3 位元流），該已編碼位元流並包含音頻資料片段（例如，在圖 4 所示訊框的 AB0-AB5 片段）與元資料片段，在此該等音頻資料片段指示音頻資料，而且至少一些元資料片段的各個包括響度處理狀態元資料（LPSM），以及可選地

亦包括節目邊界元資料。LPSM 係存在於以下格式中的位元流中。包括 LPSM 之元資料片段的各個係被包括在該位元流之訊框的 BSI 片段的“adbsi”欄位中，或者在該位元流之訊框的輔助資料欄位中，或者該位元流之訊框的無用位元片段中。包括 LPSM 的各元資料片段包括一 LPSM 酬載（或容器）片段，其具有以下格式：

一標頭（基本上包括至少一個識別值，例如，LPSM 格式版本、長度、時期、數目與子串流相關值，其係在下文表 2 中指出）；以及

在該標頭以後，LPSM 與亦可選地節目邊界元資料。該節目邊界元資料包括一節目邊界訊框數目，以及一代碼值（例如，一“偏移_存在”值，其指示該訊框是否僅包括一節目邊界訊框數或者一節目邊界訊框數與一偏移值兩者，以及（在一些情形中）一偏移值。LPSM 包括：

至少一個對話指示值（例如，表 2 的參數“對話頻道”），其指示對應的音頻資料是否指示對話或不指示對話（例如，哪些對應音頻資料的頻道指示對話）。該對話指示值可指示對話是否存在於該對應音頻資料之頻道的任一組合或全部中；

至少一個響度規則遵守值（例如，表 2 的參數“響度規則類型”），指示對應的音頻資料是否遵守一指定組響度規則；

至少一個響度處理值（例如，表 2 之參數“對話閘控響度校正旗標”、“響度校正類型”的其中一或多個），

指示至少一種響度處理，其已經施行在該對應的音頻資料上；以及

至少一個響度值（例如，表 2 參數“ITU 相關閘控響度”、“ITU 演講閘控響度”、“ITU (EBU 3341) 短期 3s 響度”與“真實峰值”的其中一個或多個）指示該對應音頻資料的至少一個響度（例如，峰值或平均響度）特徵。

在打算、使用或產生指示對應音頻資料之至少一個響度值之本發明的任一項實施例中，該響度值可指示被使用來處理該音頻資料之響度及/或動態範圍的至少一個響度測量特徵。

在一些實施程序中，在位元流之訊框之“adbsi”欄位、或輔助資料欄位或無用位元片段中之元資料片段的各個具有以下格式：

一核心標頭（基本上包括識別該元資料片段之開頭的一同步字，其後接著識別值，例如，核心元件版本、長度、與時期、延伸元件數目與子串流相關值，其係在下文表 1 中指出）；以及

在該核心標頭以後，至少一個保護值（例如，HMAC 摘要與音頻指紋值，在此 HMAC 摘要為在一整個訊框之音頻資料、核心元件與全部延伸元件上所計算的 256-位元 HMAC 摘要（digest）（使用 SHA-2 演算法），如表 1 所指示），其對至少一個響度處理狀態元資料或者該對應音頻資料之解密、鑑定或驗證的至少其中一個有用；以及

亦在該核心標頭以後，假如該元資料片段包括 LPSM、LPSM 酬載識別（“ID”）與 LPSM 酬載尺寸值，其將接著的元資料識別為 LPSM 酬載並且指示 LPSM 酬載的尺寸。LPSM 酬載片段（較佳地具有上文具體指定的格式）接著 LPSM 酬載 ID 與 LPSM 酬載尺寸值。

在先前篇幅中所說明之該類型的一些實施例中，在該訊框之輔助資料欄位（或“addbsi”欄位或無用位元欄位）中之元資料片段的各個具有三階結構：

一高階結構，包括指示該輔助資料（或者 addbsi）欄位是否包括元資料的一旗標、指示哪（些）種元資料存在的至少一個 ID 值、以及基本上亦包括指示（例如，各類型的）元資料中有多少位元存在（假如元資料存在）的一值。一種可存在的元資料為 LSPM，另一種可存在的元資料為節目邊界元資料，且另一種可存在的元資料為媒體研究元資料；

一中階結構，包含各識別類型元資料的核心元件（例如，各識別類型元資料之上文所描述類型的核心標頭、保護值與酬載 ID 與酬載尺寸值）；以及

一低階結構，包含一個核心元件的各酬載（例如，LPSM 酬載，假如一個酬載係藉由核心元件被識別為存在，及/或另一類型的元資料酬載，假如一個被核心元件識別為存在）。

在此類三階結構中的資料值可被套入。例如，LPSM 酬載及/或由核心元件所識別之另一元資料酬載的保護

值，其可在該核心元件所識別的各酬載以後被包括（以及因此在該核心元件的核心標頭以後）。在一項實例中，核心標頭可識別 LPSM 酬載與另一元資料酬載，第一酬載（例如，LPSM 酬載）的酬載 ID 與酬載尺寸值可接著該核心標頭，第一酬載本身可接著該 ID 與尺寸值，第二酬載的酬載 ID 與酬載尺寸值可接著第一酬載，第二酬載本身可接著這些 ID 與尺寸值，且該等酬載之任一個或兩個的保護值（或者用於核心元件值與該等酬載的任一個或兩個）可接著最後酬載。

在一些實施例中，在一訊框之輔助資料欄位中（或者“addbsi”欄位或無用位元片段）之元資料片段的核心理件包含一核心標頭（基本上包括識別值，例如核心元件版本），而且在該核心標頭以後：指示指紋資料是否被包括以用於該元資料片段之元資料的值、指示外部資料（與對應該元資料片段之元資料的音頻資料相關）是否存在的值、由該核心元件所識別之各類型元資料（例如，LPSM、及/或除了 LPSM 以外之一類型的元資料）的酬載 ID 與酬載尺寸值、與由該核心元件所識別之至少一類型元資料的保護值。該元資料片段的元資料酬載接著該核心標頭，並且（在一些情形中）被套入核心元件值內。

在另一較佳格式中，已編碼位元流為杜比 E 位元流，而且包括 LPSM 之元資料片段的各個（以及可選地亦包括節目邊界元資料）係被包括在杜比 E 保護帶間隔的第一 N 樣本位置中。

在另一種類的實施例中，本發明為被耦合與組態以接收包含音頻資料片段與元資料片段之已編碼音頻位元流的 APU（例如，解碼器），在此，該等音頻資料片段指示音頻資料，而且至少一些元資料片段的各個包括響度處理狀態元資料（LPSM）以及可選地亦包括節目邊界元資料，以及從該位元流擷取 LPSM，產生已解碼的音頻資料，以回應該音頻資料，以及使用 LPSM 施行至少一個自適應響度處理操作於音頻資料上。在此種類中的一些實施例亦包括耦合到 APU 的後處理器，其中該後處理器被耦合且組態以使用 LPSM 施行至少一個自適應響度處理操作於音頻資料上。

在另一種類的實施例中，本發明為音頻處理單元（APU），其包括一緩衝器記憶體（緩衝器）與耦合到該緩衝器的一處理子系統，其中 APU 係被耦合以接收一已編碼音頻位元流，其包含音頻資料片段與元資料片段，在此，該等音頻資料片段指示出音頻資料，而且至少一些元資料片段的各個包括響度處理狀態元資料（LPSM）以及可選地亦包括節目邊界元資料，該緩衝器儲存（例如，以非過渡方式）該已編碼音頻位元流的至少一個訊框，而且該處理子系統係被組態以從該位元流擷取 LPSM，以及使用 LPSM 施行至少一個自適應響度處理操作於音頻資料上。在此種類中的典型實施例，APU 為編碼器、解碼器與後處理器的其中一種。

在本發明方法的一些實施程序中，所產生的音頻位元

流為 AC-3 已編碼位元流、E-AC-3 位元流或杜比 E 位元流的其中一個，其包括響度處理狀態元資料以及其他元資料（例如，DIALNORM 元資料參數、動態範圍控制元資料參數與其他元資料參數）。在本方法的一些其他實施程序中，所產生的音頻位元流為另一類型的已編碼位元流。

本發明的態樣包括被組態（例如，被程式化）以施行本發明方法之任何實施例的系統或裝置，以及儲存用來實施本發明方法或其步驟之任何實施例之代碼（例如，以非過渡方式）的電腦可讀取媒體（例如，碟片）。例如，本發明系統為或包括可程式化一般目的處理器、數位訊號處理器或微處理器，其係以軟體或韌體來程式化及/或另外被組態以施行許多操作之任一個於資料上，包括本發明方法或其步驟的實施例。此一般目的處理器為或包括一電腦系統，其包括被程式化（及/或被組態）以回應提示（assert）至其的資料而施行本發明方法（或其步驟）之實施例的一輸入裝置、一記憶體與處理電路。

[標誌與術語]

在此揭露之全文，包括在申請專利範圍中，施行一操作於一訊號或資料“上”（例如，過濾、定標、變換、或施加增益到該訊號或資料）的表達式，其係在廣泛的意義中，被使用來指示直接施行該操作於該訊號或資料上，或者在該訊號或資料的處理版本上（例如，在執行操作於其上以前已經受到初步過濾或預處理之訊號的版本上）。

在包括在申請專利範圍中的整個此揭露，表達式“系統”在廣泛的意義中，係被使用來指示裝置、系統或子系統。例如，實施解碼器的子系統可被視為一解碼器系統，且包括此類子系統的系統（例如，回應複數個輸入來產生 X 輸出訊號的系統，其中該子系統產生 M 個輸入，且其他 X-M 輸入則從外部來源接收）亦被稱為解碼器系統。

此揭露之全文，包括在申請專利範圍中，在廣泛的意義中，術語“處理器”被使用來指示可程式化或另外可組態（例如，用軟體或韌體）以施行操作於資料（例如，音頻或視頻或其他影像資料）上的系統或裝置。處理器的實例包括場可程式化閘極陣列（或其他可組態積體電路或晶片組）、被程式化及/或另外被組態以施行管線處理於音頻或其他聲音資料上的數位訊號處理器、可程式化的一般目的處理器或電腦、與一可程式化微處理器晶片或晶片組。

此揭露之全文，包括在申請專利範圍中，表達式“音頻處理器”與“音頻處理單元”可交換地使用，而且在廣泛的意義中，指示被組態以處理音頻資料的一系統。音頻處理單元的實例包括但不限於編碼器（例如，轉碼器）、解碼器、編解碼器、預處理系統、後處理系統、與位元流處理系統（有時稱為位元流處理工具）。

此揭露之全文，包括在申請專利範圍中，表達式“處理狀態元資料”（例如，如在表達式“響度處理狀態元資料”中）意指與對應音頻資料分開且不同的資料（亦包括

處理狀態元資料之音頻資料流的音頻內容)。處理狀態元資料係與音頻資料有關，其指示該對應音頻資料的響度處理狀態(例如，哪種處理已經在音頻資料上施行)，並且基本上亦指示該音頻資料的至少一個特色或特徵。具有音頻資料之該處理狀態元資料的結合係與時間同步。因此，目前(最近接收或更新)的處理狀態元資料指示該對應的音頻資料同時地包含所指出音頻資料處理類型的結果。在一些情形中，處理狀態元資料包括處理歷史及/或被使用在及/或從該指示類型之處理所取得的一些或全部參數。此外，處理狀態元資料包括該對應音頻資料的至少一個特色或特徵，其係已經被計算或從該音頻資料被擷取。處理狀態元資料亦包括不相關或從該對應音頻資料的任何處理取得的其他元資料。例如，第三方資料、追蹤資訊、識別器、專屬或標準資訊、使用者註記資料、使用者偏好資料等等，可藉由特定音頻處理單元來添加，以直通到其他音頻處理單元上。

此揭露之全文，包括在申請專利範圍中，表達式“響度處理狀態元資料”(或“LPSM”)指示處理狀態元件，其指示出對應音頻資料的響度處理狀態(例如，哪類型的響度處理已經在音頻資料上施行)以及基本上亦指示該對應音頻資料的至少一個特色或特徵(例如，響度)。響度處理狀態元資料包括不是(亦即，當它被單獨考慮時)響度處理狀態元資料的資料(例如，其他元資料)。

此揭露之全文，包括在申請專利範圍中，表達式“頻

道”（或“音頻頻道”）指示一單音音頻訊號。

此揭露之全文，包括在申請專利範圍中，表達式“音頻節目”指示一組一或多個音頻頻道，以及亦可選地相關元資料（例如，說明一希望空間音頻呈現的元資料，及/或 LPSM，及/或節目邊界元資料）。

此揭露之全文，包括在該申請專利範圍中，表達式“節目邊界元資料”指示一已編碼音頻位元流的元資料，在此該已編碼音頻位元流指示至少一個音頻節目（例如，兩或多個音頻節目），且該節目邊界元資料指示在至少一個該音頻節目之至少一個邊界（開始及/或結束）之位元流中的位置。例如，（指示出音頻節目之已編碼音頻位元流的）該節目邊界元資料包括指示出該節目開始之位置（例如，該位元流之第“N”訊框的開頭，或者該位元流之第“N”訊框的第“M”樣本位置）的元資料，以及指示出該節目結束之位置（例如，該位元流之第“J”訊框的開頭，或者該位元流之第“J”訊框的第“K”樣本位置）的額外元資料。

此揭露之全文，包括在申請專利範圍中，術語“耦合”或“被耦合”使用來意指直接或間接連接。因此，假如第一裝置耦合到第二裝置，那連接可經過一直接連接，或者經過經由其他裝置與連接的間接連接。

【圖式簡單說明】

圖 1 為被組態以施行本發明方法實施例之系統實施例

的方塊圖。

圖 2 為本發明音頻處理單元實施例之編碼器的方塊圖。

圖 3 為本發明音頻處理單元之實施例之解碼器及與本發明音頻處理單元的另一實施例耦合之後處理器的方塊圖。

圖 4 為 AC-3 訊框的圖，其包括它被分成的片段。

圖 5 為 AC-3 訊框之同步資訊 (SI) 片段的圖，其包括它被分成的片段。

圖 6 為 AC-3 訊框之位元流資訊 (BSI) 片段的圖，其包括它被分成的片段。

圖 7 為 E-AC-3 訊框的圖，其包括它被分成的片段。

圖 8 為已編碼音頻位元流之訊框的圖，其包括格式係根據本發明實施例的節目邊界元資料。

圖 9 為圖 8 之已編碼音頻位元流之其他訊框的圖。這些訊框其中有一些包括具有根據本發明實施例之格式的節目邊界元資料。

圖 10 為兩個已編碼音頻位元流的圖：一位元流 (IEB)，其中一節目邊界 (標為“邊界”) 對準該位元流的兩個訊框之間的轉變，以及另一位元流 (TB)，其中一節目邊界 (標為“真實邊界”) 係藉由來自 512 樣本從該位元流之兩個訊框之間的轉變的偏移。

圖 11 為顯示四個已編碼音頻位元流的一組圖。在圖 11 頂部上的位元流 (標為“情境 1”) 指示出包括節目邊

界元資料的第一音頻節目 (P1) ，其後接著亦包括節目邊界元資料的第二音頻節目 (P2) ；第二位元流 (標為 “情境 2”) 指示出包括節目邊界元資料的第一音頻節目 (P1) ，其後接著不包括節目邊界元資料的第二音頻節目 (P2) ；第三位元流 (標為 “情境 3”) 指示出一截斷的第一音頻節目 (P1) ，其包括節目邊界元資料而且其已經接合包括節目邊界元資料的整個第二音頻節目 (P2) ；以及第四位元流 (標為 “情境 4”) ，指示出包括節目邊界元資料的一截斷第一音頻節目 (P1) 以及包括節目邊界元資料與接合一部份第一音頻節目的一截斷第二音頻節目 (P2) 。

【實施方式】

根據本發明的典型實施例，節目響度元資料的酬載，稱為響度處理狀態元資料 (“LPSM”) 與可選地亦稱為節目邊界元資料，其係被嵌入於亦包括音頻資料於其他片段 (音頻資料片段) 中之音頻位元流之元資料片段的一或多個保留欄位 (或槽) 中。基本上，該位元流之各訊框的至少一個片段包括 LPSM，以及該訊框的至少一個其他片段包括對應的音頻資料 (亦即，其響度處理狀態與響度係由 LPSM 指示的音頻資料)。在一些實施例中，LPSM 的資料量足夠小，以在不影響被分配以攜帶音頻資料的位元速率下被攜帶。

當兩或多個音頻處理單元必須遍及該處理鏈 (或者內

容生命週期)串聯地彼此運作時，在音頻資料處理鏈中的通訊響度處理狀態元資料特別有用。沒有將響度處理狀態元資料包括在音頻位元流中，譬如品質、位準與空間退化的嚴重媒體處理問題可發生，例如當兩或多個音頻編解碼器被使用於該鏈中時，且在位元流到媒體消耗裝置(或者該位元流之音頻內容的顯現點)的旅程期間內，單一端點的音量分級會被施加超過一次。

圖 1 為一例示性音頻處理鏈(一音頻資料處理系統)的方塊圖，其中該系統的一或多個元件可根據本發明實施例被組態。該系統包括以下元件，其係如所示地被耦合在一起：預處理單元、編碼器、訊號分析與元資料校正單元、轉碼器、解碼器與預處理單元。在所示系統的變化中，該等元件的其中一個或多個會被省略，或者額外的音頻資料處理單元會被包括。

在一些實施程序中，圖 1 的預處理單元係被組態，以接收包含音頻內容的 PCM(時域)樣本為輸入，以及輸出已處理的 PCM 樣本。編碼器被組態以接收 PCM 樣本為輸入並且輸出指示音頻內容的已編碼(例如，已壓縮)音頻位元流。指示音頻內容的位元流資料在本文中有時稱為“音頻資料”。假如編碼器根據本發明的典型實施例來組態，從編碼器輸出的音頻位元流包括響度處理狀態元資料(以及基本上亦包括其他元資料，可選地包括節目邊界元資料)以及音頻資料。

圖 1 的訊號分析與元資料校正單元可接收一或多個已

編碼音頻位元流為輸入，並且藉由施行訊號分析（例如，使用節目邊界元資料於已編碼的音頻位元流中），決定（例如，驗證）在各已編碼音頻位元流中的處理狀態元資料是否正確。假如該訊號分析與元資料校正單元發現已包括的元資料無效時，它基本上會以從訊號分析得到的正確值來替代該不正確的值。因此，從該訊號分析與元資料校正單所輸出的各已編碼音頻位元流，可包括已校正（或未校正）處理狀態元資料以及已編碼音頻資料。

圖 1 的轉碼器可接收已編碼的音頻位元流為輸入，以及回應地輸出已修改（例如，被不同編碼）的音頻位元流（例如，藉由以不同的編碼格式來解碼一輸入串流與重新編碼該解碼串流）。假如轉碼器根據本發明的典型實施例來組態，那麼從該轉碼器輸出的音頻位元流則包括響度處理狀態元資料（以及基本上亦包括其他元資料）以及已編碼的音頻資料。該元資料已經被包括在該位元流中。

圖 1 的解碼器可接收已編碼（例如，已壓縮）的音頻位元流為輸入，並且（回應地）輸出解碼 PCM 音頻樣本串流。假如該解碼器係根據本發明的典型實施例來組態，在典型操作中之解碼器的輸出係或包括以下任一個：

一音頻樣本串流，以及從一輸入已編碼位元流所擷取的一對應響度處理狀態元資料串流（以及基本上亦包括其他元資料）；或者

一音頻樣本串流，以及從一輸入已編碼位元流所擷取的響度處理狀態元資料（以及基本上亦包括其他元資料）

所決定之控制位元的一對應串流；或者

一音頻樣本串流，沒有處理狀態元資料或從處理狀態元資料所決定之控制位元的一對應串流。在此最後情形中，該解碼器可從該輸入編碼位元流擷取響度處理狀態元資料（及/或其他元資料）並且將至少一個操作施行於所擷取的元資料上（例如，驗證），即使它沒有輸出所擷取的元資料或者控制從那決定的位元。

藉由根據本發明的典型實施例來組態圖 1 的後處理單元，該後處理單元係被組態，以接收一解碼 PCM 音頻樣本串流，以及使用以該等樣本所接收的響度處理狀態元資料（以及典型地亦使用其他元資料）來施行後處理於上（例如，該音頻內容的音量分級），或者控制以該等樣本來接收的位元（藉由來自響度處理狀態元資料與典型亦來自其他元資料的解碼器所決定）。後處理單元基本上亦被組態以使得該後處理音頻內容由一或多個揚聲器所播放。

本發明的典型實施例提供一加強的音頻處理鏈，其中音頻處理單元（例如，編碼器、解碼器、轉碼器以及預與後處理單元）根據由該音頻處理單元各自接收的響度處理狀態元資料所指示之元資料的同期狀態，來修改它們各自的處理以施加到音頻資料。

輸入到圖 1 系統之任何音頻處理單元（例如，圖 1 的編碼器或轉碼器）的音頻資料可包括響度處理狀態元資料（以及可選地亦包括其他元資料）以及音頻資料（例如，已編碼音頻資料）。此元資料已經藉由根據本發明實施例

之圖 1 系統的另一元件（或者另一來源，沒有顯示於圖 1 中）被包括在輸入音頻中。接收（具有元資料）輸入音頻的處理單元可被組態以施行至少一個操作（例如，驗證）於元資料上，或者回應該元資料（例如，該輸入音頻的自適應處理），以及典型地亦將該元資料、該元資料的處理版本或者從該元資料所決定的控制位元包括在它的輸出音頻中。

本發明音頻處理單元（或音頻處理器）的典型實施例係被組態，以依據由對應該音頻資料的響度處理狀態元資料所指示之音頻資料的狀態來施行音頻資料的自適應處理。在一些實施例中，自適應處理為（或包括）響度處理（假如該元資料指示該響度處理或與之類似的處理尚未被施行於音頻資料上），但不是（且不包括）響度處理（假如該元資料指示此響度處理或與之類似的處理已經被施行在音頻資料上）。在一些實施例中，該自適應處理為或包括元資料驗證（例如，在元資料驗證子單元中施行），以確保該音頻處理單元，以響度處理狀態元資料所指示之音頻資料的狀態為基礎，來施行該音頻資料的其他自適應處理。在一些實施例中，該驗證決定與該音頻資料有關（例如，被包括在位元流中）之響度處理狀態元資料的可靠度。例如，假如該元資料被驗證為可靠，那麼來自一種先前施行音頻處理的結果則可被再度使用，且相同類型之音頻處理的新性能則可被避免。另一方面，假如該元資料被發現已經被竄改（或者是不可靠），那麼故意在先前施行

的該類型元資料（如不可靠元資料所指示）可由音頻處理單元所重複，及/或其他處理可藉由音頻處理單元被施行在該元資料及/或該音頻資料上。假如該單元決定處理狀態元資料有效（例如，以所擷取之加密值與參考加密值之匹配為基礎），該音頻處理單元亦可被組態，以將響度處理狀態元資料（例如，存在於媒體位元流中）為正確的訊號發到在增強媒體處理鏈中之下游的其他音頻處理單元。

圖 2 為本發明音頻處理單元實施例之編碼器（100）的方塊圖。編碼器 100 之組件或元件的任一者，其係以在硬體、軟體或硬體及軟體之組合中的一或多個處理及/或一或多個電路來實施（例如，特定應用積體電路（ASIC）、現場可程式閘陣列（FPGA）或其他積體電路）。編碼器 100 包含訊框緩衝器 110、剖析器 111、解碼器 101、音頻狀態驗證器 102、響度處理台 103、音頻流選擇台 104、編碼器 105、填充器/格式器台 107、元資料產生台 106、對話響度測量子系統 108 與訊框緩衝器 109，其係如所示地連接。基本上，編碼器 100 亦包括其他處理元件（未顯示）。

編碼器 100（其為轉碼器）係被組態以將輸入音頻位元流（例如，其為 AC-3 位元流、E-AC-3 位元流或杜比 E 位元流的其中一者）轉換成已編碼輸出音頻位元流（其例如為 AC-3 位元流、E-AC-3 位元流或杜比 E 位元流的另一者），其包括藉由使用被包括在輸入位元流中的響度處理狀態元資料來施行自適應與自動化響度處理。例如，編碼

器 100 可被組態，以將一輸入杜比 E 位元流（基本上用於生產與廣播設備中，但非用於接收廣播至其之音頻節目的消費者裝置中的格式）轉換為 AC-3 或 E-AC-3 格式的已編碼輸出音頻位元流（適合廣播到消費者裝置）。

圖 2 的系統亦包括已編碼音頻傳送子系統 150（其儲存及/或傳送從編碼器 100 輸出的已編碼位元流）與解碼器 152。從編碼器 100 輸出的已編碼音頻位元流可藉由子系統 150 所儲存（例如，呈 DVD 或藍芽光碟形式），或者藉由子系統 150 傳送（其係可實施一傳送鏈或網路），或者可藉由子系統 150 來儲存與傳送兩者。解碼器 152 係被組態以解碼它經由子系統 150 所接收的已編碼音頻位元流（由編碼器 100 產生），包括藉由從該位元流的各訊框擷取響度處理狀態元資料（LPSM）（以及可選地亦包括從該位元流擷取節目邊界元資料），以及產生解碼音頻資料。基本上，解碼器 152 係被組態以使用 LPSM（以及可選地亦使用節目邊界元資料）來施行自適應響度處理於已解碼音頻資料上，及/或轉送已解碼音頻資料與 LPSM 到後處理器，其係被組態以使用 LPSM（以及可選地亦使用節目邊界元資料）來施行自適應響度處理於已解碼音頻資料上。基本上，解碼器 152 包括將從子系統 150 接收之已編碼音頻位元流儲存（例如，呈非過渡方式）的緩衝器。

編碼器 100 與解碼器 152 的各種實施程序係被組態以施行本發明方法的不同實施例。

訊框緩衝器 110 為被耦合以接收已編碼輸入音頻位元

流的緩衝器記憶體。在操作時，緩衝器 110 儲存（例如，以非過渡方式）已編碼音頻位元流的至少一個訊框，而且該已編碼音頻位元流的一序列訊框係從緩衝器 110 被提示（assert）到剖析器 111。

剖析器 111 係被耦合與組態以自已編碼輸入音頻之各訊框擷取響度處理狀態元資料（LPSM）及可選地節目邊界元資料（及/或其他元資料），其中該已編碼輸入音頻中包括此等之資料，以提示至少該 LPSM（以及可選地顯示節目邊界元資料及/或其他元資料）到音頻狀態驗證器 102、響度處理台 103、台 106 與子系統 108，以從已編碼輸入音頻擷取音頻資料，並且提示該音頻資料到解碼器 101。編碼器 100 的解碼器 101 係被組態以解碼該音頻資料以產生已解碼的音頻資料，並且提示已解碼的音頻資料到響度處理台 103、音頻流選擇台 104、子系統 108 以及基本上亦提示到狀態驗證器 102。

狀態驗證器 102 係被組態以鑑定與驗證被提示至其的 LPSM（以及可選地其他元資料）。在一些實施例中，LPSM 為（或被包括在）已經被包括在輸入位元流中的資料方塊（例如，根據本發明實施例）。該方塊包含加密散列（以散列為基礎的訊息鑑定碼或“HMAC”），以用來處理 LPSM（以及可選地亦處理其他元資料）及/或底層音頻資料（從解碼器 101 提供到驗證器 102）。該資料方塊可在這些實施例中被數位簽署，以致於下游音頻處理單元能夠相當簡單地鑑定與驗證該處理狀態元資料。

例如，HMAC 被使用來產生一摘要，且包括在本發明位元流中的保護值包括該摘要。該摘要可如下地產生，以用於 AC-3 訊框：

- 1.在 AC-3 資料與 LPSM 被編碼以後，訊框資料位元組（連接訊框_資料 # 1 與訊框_資料 # 2）與 LPSM 資料位元組被使用當作散列函數 HMAC 的輸入。存在於輔助資料欄位裡面的其他資料不會被考慮用來計算該摘要。此其他資料為既不屬於 AC-3 資料也不屬於 LPSM 資料的位元組。被包括在 LPSM 中的保護位元不被考慮用來計算 HMAC 摘要。

- 2.在該摘要被計算以後，它可被寫入到保留用於保護位元之欄位中的位元流內。

- 3.產生完整 AC-3 訊框的最後步驟為 CRC-檢查的計算。這在該訊框的非常末期被寫入，而且屬於此訊框的全部資料係被考慮在內，其包括 LPSM 位元。

包括但不限於一或多個非 HMAC 加密方法之任一個的其他加密方法可被使用於驗證 LPSM（例如，在驗證器 102 中），以確保 LPSM 及/或底層音頻資料的安全傳送與接收。例如，（使用此類加密方法）驗證可被施行於各音頻處理單元中，該單元接收本發明音頻位元流的實施例，以決定是否被包括在該位元流中的響度處理狀態元資料與對應音頻資料已經受到（及/或已經起因於）具體響度處理（如元資料所指示）並且在執行此具體響度處理以後已經不會被修改。

狀態驗證器 102 提示控制資料到音頻串流選擇台 104、元資料產生器 106 與對話響度測量系統 108，以指示該驗證操作的結果。回應該控制資料，台 104 可選擇任一者（並且直通到編碼器 105）：

響度處理台 103 的自適應處理輸出（例如，當 LPSM 指示來自解碼器 101 的音頻資料輸出不會受到具體類型的響度處理，且來自驗證器 120 的控制位元指示 LPSM 為正確）；或者

來自解碼器 101 的音頻資料輸出（例如，當 LPSM 指示出來自解碼器 101 的音頻資料輸出已經受到由台 103 所施行之具體類型的響度處理，且來自驗證器 102 的控制位元指示 LPSM 為正確）。

編碼器 100 的台 103 係被組態以施行自適應響度處理於從解碼器 101 輸出的解碼音頻資料上，其係以由解碼器 101 所擷取之 LPSM 指示的一或多個音頻資料特徵為基礎。台 103 為自適應轉換域即時響度與動態範圍控制處理器。台 103 可接收使用者輸入（例如，使用者目標響度/動態範圍值或對白歸一化值）或其他元資料輸入（例如，一或多個類型的第三方資料、追蹤資訊、識別器、專屬或標準資訊、使用者注解資料、使用者偏好資料等等）及/或其他輸入（例如，來自印指紋製程）及使用此輸入以處理從解碼器 101 輸出的解碼音頻資料。台 130 可施行自適應響度處理於指示出單一音頻節目（如剖析器 111 所擷取之節目邊界元資料所指示）的解碼輸出資料上（來自解碼

器 101 的輸出)，並且可回應接收解碼音頻資料（來自解碼器 101 的輸出）來重設該響度處理，其指示由剖析器 111 所擷取之節目邊界元資料所指示的不同音頻節目。

對話響度測量子系統 108 可操作，以決定解碼音頻（來自解碼器 101）之片段的響度，其指示例如使用解碼器 101 所擷取之 LPSM（及/或其他元資料）的對話（或其他演講），當來自驗證器 120 的控制位元指示 LPSM 無效時。當來自驗證器 102 之控制位元指示 LPSM 正確時，當 LPSM 指示（來自解碼器 101）已解碼音頻之對話（或其他演講）片段的先前決定響度時，對話響度測量子系統 108 的操作可被失能。子系統 108 可施行響度測量於指示出訊號音頻節目的解碼音頻資料上（如藉由剖析器 111 所擷取之節目邊界元資料所指示），並且可回應接收指示出此節目邊界元資料所指示之不同音頻節目的解碼音頻資料來重新設定該測量。

有用工具（例如，杜比 LM100 響度計）存在為了方便且輕易地測量在音頻內容中的對話位準。本發明 APU 的一些實施例（例如，編碼器 100 的台 108）係被實施以包括此類工具（或施行其功能），以測量音頻位元流之音頻內容的平均對話響度（例如，自編碼器 100 之解碼器 101 被提示到台 108 的解碼 AC-3 位元流）。

假如台 108 係被實施以測量音頻資料的真實平均對話響度，該測量可包括一種將主要包含演講之音頻內容片段隔開的步驟。主要為演講的該音頻片段隨後根據響度測量

演算法來處理。就從 AC-3 位元流解碼的音頻資料而言，此演算法為標準的 K-加權響度測量（根據國際標準 ITU-R BS.1770）。或者，其他響度測量可被使用（例如，那些以響度之心理聲學模組為基礎者）。

演講片段的隔離對測量音頻資料的平均對話響度不重要。不過，它改善測量的準確性而且基本上提供來自聆聽者觀點的更多令人滿意的結果。因為並非全部音頻內容包含對話（演講），所以全部音頻內容的響度測量則可提供該音頻之對話位準的充分近似，而使演講存在。

元資料產生器 106 產生（及/或直通到台 107）由台 107 所包括的元資料於從編碼器 100 所輸出的已編碼位元流中。元資料產生器 106 可將編碼器 101 及/或剖析器 111 所擷取的 LPSM（以及可選地亦將節目邊界元資料及/或其他元資料）直通到台 107（例如，當來自驗證器 102 的控制位元指示 LPSM 及/或其他元資料是正確時）或者產生新的 LPSM（以及可選地亦產生節目邊界元資料及/或其他元資料）與將新的元資料提示到台 107（例如，當來自驗證器 102 的控制位元指示由解碼器 101 所擷取的 LPSM 及/或其他元資料無效時），或者它可將由解碼器 101 及/或剖析器 111 所擷取之元資料與最新產生的元資料的組合提示到台 107。元資料產生器 106 可包括由子系統 108 所產生的響度資料、以及至少一個指示由子系統 108 所施行的響度處理類型的值）於 LPSM 中該 LPSM 被提示至台 107，以用來包括在欲從編碼器 100 輸出的已編碼位元流

中。

元資料產生器 106 可產生保護位元（其係由以散列為基礎的訊息鑑定碼或“HMAC”所組成或包括），其對被包括在已編碼位元流之 LPSM（以及可選地亦對其他元資料）及/或被包括在已編碼位元流之底層音頻資料之解密、鑑定或驗證的至少一個有用。元資料產生器 106 可提供此等保護位元到台 107，以用於包括在已編碼位元流中。

在典型的操作中，對話響度測量子系統 108 處理從解碼器 101 輸出的音頻資料，以回應地產生響度值（例如，開控與未開控對話響度值）與動態範圍值。回應這些值，元資料產生器 106 可產生響度處理狀態元資料（LPSM），以用來包括（藉由填充器/格式器 107）到已編碼位元流內，以從編碼器 100 輸出。

此外，可選地或者替代地，編碼器 100 之 106 及/或 108 的子系統可施行音頻資料的額外分析，以產生指示該音頻資料之至少一個特徵的元資料，以用來包括在從台 107 輸出的已編碼位元流。

編碼器 105 編碼（例如，藉由施行壓縮於上）從選擇台 104 輸出的音頻資料，並且將已編碼音頻提示到台 107，以用來包括在從台 107 輸出的已編碼位元流。

台 107 將來自編碼器 105 的已編碼音頻與來自產生器 106 的元資料（包括 LPSM）多工化，以產生已編碼位元流，以從台 107 輸出，較佳地以致於該已編碼位元流具有

由本發明較佳實施例所具體指定的格式。

訊框緩衝器 109 為一緩衝器記憶體，其儲存（例如，以非過渡方式）從台 107 輸出之已編碼音頻位元流的至少一個訊框，且已編碼音頻位元流的一序列訊框隨後從緩衝器 109 被當作編碼器 100 之輸出地被提示到傳送系統 150。

元資料產生器 106 所產生並且藉由台 107 被包括在已編碼位元流中的 LPSM，其係指示對應音頻資料的響度處理狀態（例如，哪種類型的響度處理已經施行於音頻資料上）與對應音頻資料的響度（例如，已測量對話響度、聞控及/或未聞控響度、及/或動態範圍）。

在本文中，在音頻資料上施行的響度及/或位準測量的“聞控”意指一具體位準或響度臨界值，在此，超過該臨界值的計算值係被包括在最後的測量中（例如，忽略在最後測量值中低於-60dBFS 的短期響度值）。在絕對值上的聞控意指一固定位準或響度，然而，在相對值上的聞控意指取決於目前“未聞控”測量值的值。

在編碼器 100 的一些實施程序中，在記憶體 109 中被緩衝（以及輸出到傳送系統 150）的已編碼位元流為 AC-3 位元流或 E-AC-3 位元流，並且包含音頻資料片段（例如在圖 4 所示之訊框的 AB0-AB5 片段）及元資料片段，在此該等音頻資料片段指示音頻資料，且至少一些元資料片段的各個包括響度處理狀態元資料（LPSM）。台 107 用以下的格式將 LPSM（以及可選地亦將節目邊界元資料）

插入到位元流內。包括 LPSM（以及可選地亦包括節目邊界元資料）的各元資料片段係被包括在該位元流的無用位元片段中（例如，如圖 4 或圖 7 所示的無用位元片段“W”）或者該位元流之一訊框之位元流資訊（“BSI”）片段的“addbsi”欄位，或者在該位元流之一訊框結束時的輔助資料欄位中（例如，在圖 4 或圖 7 中所示的 AUX 片段）。該位元流之一訊框包括一或兩個元資料片段，其中各個包括 LPSM，且假如該訊框包括兩個元資料片段，一個存在於該訊框的 addbsi 欄位中，且另一個在該訊框的 AUX 欄位中。在一些實施例中，包括 LPSM 的各元資料片段包括具有以下格式的 LPSM 酬載（或容器）片段：

一標頭，（基本上包括一同步字，其識別 LPSM 酬載的開頭，其後接著至少一個識別值，例如，LPSM 格式版本、長度、時期、數目與子串流相關值，如在下文表 2 中所指示），以及

在該標頭以後，

至少一個對話指示值（例如，表 2 的參數“對話頻道”），其指示對應的音頻資料是否指示對話或不指示對話（例如，哪些對應音頻資料的頻道指示對話）；

至少一個響度規則遵守值（例如，表 2 的參數“響度規則類型”），其指示對應的音頻資料是否遵守一指定組響度規則；

至少一個響度處理值（例如，表 2 參數“對話閘控響

度校正旗標”、“響度校正類型”的其中一個或多個），其指示已經在該對應音頻資料上施行的至少一種響度處理；以及

至少一個響度值（例如，表 2 之參數“ITU 相關閘控響度”、“ITU 演講閘控響度”、“ITU (EBU 3341) 短期 3s 響度”與“真實峰值”的其中一個或多個），指示該對應音頻資料的至少一個響度（例如，峰值或平均響度）特徵。

在一些實施例中，包含 LPSM 與節目邊界元資料的各元資料片段包含一核心標頭（以及可選地亦包含額外核心元件），以及在該核心標頭（或者該核心標頭與其他核心元件）以後，LPSM 酬載（或容器）片段具有以下格式：

一標頭，基本上包括至少一個識別值（例如，LPSM 格式版本、長度、時期、數目與子串流相關值，如在本文所陳述之表 2 中所指出），以及

在該標頭以後，LPSM 與節目邊界元資料。該節目邊界元資料包括一節目邊界訊框數目，以及一代碼值（例如，一“偏移_存在”值），其指示該訊框是否僅包括一節目邊界訊框數或者一節目邊界訊框數與一偏移值兩者，以及（在一些情形中）一偏移值。

在一些實施程序中，藉由台 107 插入到該位元流之一訊框之無用位元片段或“額外位元流 (adbsi)”欄位或輔助資料欄位內的各元資料片段具有以下格式：

一核心標頭（基本上包括識別該元資料片段之開頭的

一同步字，其後接著識別值，例如，核心元件版本、長度、與時期、延伸元件數目與子串流相關值，其係在下文表 1 中指出）；以及

在該核心標頭以後，至少一個保護值（例如，表 1 的 HMAC 摘要與音頻指紋值），其對響度處理狀態元資料或者該對應音頻資料之至少一個之解密、鑑定或驗證的至少一個有用；以及

亦在該核心標頭以後，假如該元資料片段包括 LPSM、LPSM 酬載識別（“ID”）與 LPSM 酬載尺寸值，其係將以下的元資料識別為 LPSM 酬載並且指示 LPSM 酬載的尺寸。

LPSM 酬載（或容器）片段（較佳地具有上文具體指定的格式）接著 LPSM 酬載 ID 與 LPSM 酬載尺寸值。

在一些實施例中，在一訊框之輔助資料欄位（或“addbsi”欄位）中之元資料片段的各個具有三階結構：

一高階結構，包括指示該輔助資料（或者額外位元流）欄位是否包括元資料的一旗標、指示哪（些）種元資料存在的至少一個 ID 值、以及基本上亦包括指示（例如，各類型的）元資料中有多少位元存在（假如元資料存在）的一值。可存在的一類型元資料為 LSPM，可存在的另一類型元資料為節目邊界元資料，且可存在的另一類型元資料為媒體研究元資料（例如，Nielsen 媒體研究元資料）；

一中階結構，包含各識別類型元資料的核心元件（例

如，如上文所描述，各識別類型元資料的核心標頭、保護值與 LPSM 酬載 ID 與 LPSM 酬載尺寸值）；以及

一低階結構，包含一個核心元件的各酬載（例如，LPSM 酬載，假如一個酬載係藉由核心元件被識別為存在，及/或另一類型的元資料酬載，假如一個酬載被核心元件識別為存在）。

在此類三階結構中的資料值可被套入。例如，LPSM 酬載及/或由核心元件所識別之另一元資料酬載的保護值，其係可在該核心元件所識別的各酬載以後被包括（以及因此在該核心元件的核心標頭以後）。在一項實例中，核心標頭可識別 LPSM 酬載與另一元資料酬載，第一酬載（例如，LPSM 酬載）的酬載 ID 與酬載尺寸值可接著該核心標頭，第一酬載本身可接著該 ID 與尺寸值，第二酬載的酬載 ID 與酬載尺寸值可接著第一酬載，第二酬載本身可接著這些 ID 與尺寸值，且兩酬載的保護值（或者用於核心元件值與兩酬載）可接著最後酬載。

在一些實施例中，假如解碼器 101 接收以加密散列之根據本發明實施例來產生的音頻位元流，該解碼器係被組態以剖析並且擷取來自從該位元流所決定之資料方塊的加密散列，該方塊包含響度處理狀態元資料（LPSM）及可選地亦包括節目邊界元資料。驗證器 102 使用該加密散列來驗證所接收的位元流及/或相關元資料。例如，假如以參考加密散列與從該資料方塊所擷取的加密散列為基礎，驗證器 102 發現 LPSM 正確，那麼它可使對應音頻資料上

之處理器 103 的操作失能並且導致選擇台 104（未改變地）直通該音頻資料。此外，可選地或替代地，其他類型的加密技術可被使用，以取代以加密散列為基礎的方法。

圖 2 的編碼器 100 可決定（回應由解碼器 101 所擷取的 LPSM，以及可選地亦回應節目邊界元資料）後/預處理單元已經施行一類型的響度處理於（在元件 105、106 與 107 中）欲編碼的音頻資料上，而且（在產生器 106 中）因此可產生響度處理狀態元資料，其包括使用於先前施行響度處理中及/或從其取得的具體參數。在一些實施程序中，編碼器 100 可產生（及包括在自其輸出的已編碼位元流中）處理狀態元資料，其指示在音頻內容上的處理歷史，只要該編碼器知道已在該音頻內容上施行的該等處理類型。

圖 3 為本發明音頻處理單元之實施例之解碼器（200）以及與之耦合之後處理器（300）的方塊圖。後處理器（300）亦為本發明音頻處理單元的實施例。解碼器 200 與後處理器 300 之組件或元件的任一個係以硬體、軟體或硬體及軟體之組合中的一或多個處理及/或一或多個電路（例如，ASIC、FPGA 或其他積體電路）來實施。解碼器 200 包含訊框緩衝器 201、剖析器 205、音頻解碼器 202、音頻狀態驗證台（驗證器）203 與控制位元產生台 204，其係如所示地連接。基本上，解碼器 200 包括其他處理元件（未顯示）。

訊框緩衝器 201（緩衝記憶體）儲存（例如，以非過

渡的方式)由解碼器 200 所接收的已編碼音頻位元流的至少一個訊框。已編碼音頻位元流的一序列訊框係從緩衝器 201 被提示到剖析器 205。

剖析器 205 係被耦合與組態，以擷取響度處理狀態元資料 (LPSM) 及可選地亦擷取節目邊界元資料，以及來自己編碼輸入音頻之各訊框的其他元資料；以將至少 LPSM (以及節目邊界元資料，假如有任何被擷取的話) 提示到音頻狀態驗證器 203 與台 204；以提示 LPSM (以及可選地節目邊界元資料) 做為輸出 (例如，至後處理器 300)；以擷取來自己編碼輸入音頻的音頻資料；以及將所擷取的音頻資料提示到解碼器 202。

輸入到解碼器 200 的已編碼音頻位元流為 AC-3 位元流、E-AC-3 位元流或杜比 E 位元流的其中一個。

圖 3 的系統亦包括後處理器 300。後處理器 300 包含訊框緩衝器 301 與其他處理元件 (未顯示)，其包括被耦合到緩衝器 301 的至少一個處理元件。訊框緩衝器 301 儲存 (例如，以非過渡方式) 來自解碼器 200 之由後處理器 300 所接收之至少一個訊框的解碼音頻位元流。後處理器 300 的處理元件係被耦合且組態以接收且自適應處理從緩衝器 301 輸出之解碼音頻位元流的一序列訊框，其係使用從解碼器 202 輸出的元資料 (包括 LPSM 值) 及/或從解碼器 200 之台 204 輸出的控制位元。基本上，後處理器 300 係被組態以使用 LPSM 值與可選地亦使用節目邊界元資料來施行自適應響度處理於該解碼音頻資料上 (例如，

以用於指示單一音頻節目之音頻資料用的 LPSM 所指示的響度處理狀態及/或一或多個音頻資料特徵為基礎)。

解碼器 200 與後處理器 300 的各類實施程序係被組態以施行本發明方法的不同實施例。

解碼器 200 的音頻解碼器 202 係被組態以解碼藉由剖析器 205 所擷取的音頻資料，以產生解碼音頻資料，並且提示該解碼音頻資料當作輸出（例如到後處理器 300）。

狀態驗證器 203 係被組態以鑑定與驗證被提示至其的 LPSM（以及可選地其他元資料）。在一些實施例中，LPSM 為（或被包括在）已經被包括在輸入位元流中的資料方塊（例如，根據本發明實施例）。該方塊包含加密散列（以散列為基礎的訊息鑑定碼或“HMAC”），以用來處理 LPSM（以及可選地亦處理其他元資料）及/或底層音頻資料（從剖析器 205 及/或解碼器 202 提供到驗證器 203）。該資料方塊可在這些實施例中被數位簽署，以致於下游音頻處理單元能夠相當簡單地鑑定與驗證該處理狀態元資料。

包括但不限於一或多個非 HMAC 加密方法之任一個的其他加密方法可被使用於驗證 LPSM（例如，在驗證器 203 中），以確保 LPSM 及/或底層音頻資料的安全傳送與接收。例如，（使用此類加密方法）的驗證可被施行於各音頻處理單元中，該單元接收本發明音頻位元流的實施例，以決定是否被包括在該位元流中的響度處理狀態元資料與對應音頻資料已經受到（及/或已經起因於）具體響

度處理（如元資料所指示）並且在執行此具體響度處理以後已經不會被修改。

狀態驗證器 203 提示控制資料到控制位元產生器 204 及 / 或顯示作為輸出的控制資料（例如，到後處理器 300），以指示該驗證操作的結果。回應該控制資料（以及可選地亦回應從該輸入位元流擷取的其他元資料），台 204 可產生（並且提示到後處理器 300）任一個：

控制位元，指示從解碼器 202 輸出的解碼音頻資料已受到特定類型的響度處理（當 LPSM 指示從解碼器 202 輸出的音頻資料已經受到特定類型的響度處理，且來自驗證器 203 的控制位元指示 LPSM 是有效時）；或者

控制位元，指示從解碼器 202 輸出的解碼音頻資料應該接受特定類型的響度處理（例如，當 LPSM 指示從解碼器 202 輸出的音頻資料尚未受到特定類型的響度處理時，或者當 LPSM 指示從解碼器 202 輸出的音頻資料已經受到特定類型的響度處理，但來自驗證器 203 的控制位元卻指示 LPSM 無效時）。

或者，解碼器 200 將藉由解碼器 202 從輸入位元流所擷取的元資料以及藉由剖析器 205 從輸入位元流所擷取的 LPSM（以及可選地亦擷取節目邊界元資料）提示到後處理器 300，且後處理器 300 使用 LPSM（以及可選地亦使用節目邊界元資料）施行響度處理於解碼音頻資料上，或施行 LPSM 的驗證，而且假如該驗證指示 LPSM 有效時，隨後則使用 LPSM（以及可選地亦使用響度節目元資料）

施行響度處理於解碼音頻資料上。

在一些實施例中，假如解碼器 200 接收用加密散列之根據本發明實施例所產生的音頻位元流，解碼器係被組態以剖析且擷取來自從該位元流所決定之資料方塊的加密散列，該方塊包含響度處理狀態元資料（LPSM）。驗證器 203 可使用加密散列來驗證所接收的位元流及/或相關元資料。例如，假如以參考加密散列與從該資料方塊擷取之加密散列之間的匹配為基礎，驗證器 203 發現 LPSM 有效，那麼它可發訊號到下游的音頻處理單元（例如，後處理器 300，其為或包括音量分級單元）以（未改變地）直通該位元流的音頻資料。此外，可選地或替代地，其他類型的加密技術可被使用，以取代以加密散列為基礎的方法。

在解碼器 200 的一些實施程序中，所接收的（以及在記憶體 201 中被緩衝的）已編碼位元流為 AC-3 位元流或 E-AC-3 位元流，並且包含音頻資料片段（例如在圖 4 所示訊框的 AB0-AB5 片段）及元資料片段，在此該等音頻資料片段指示音頻資料，且至少一些元資料片段的各個包括響度處理狀態元資料（LPSM）與可選地亦包括節目邊界元資料。解碼器台 202（及/或剖析器 205）係被組態以從具有以下格式的位元流 LPSM（以及可選地亦從節目邊界元資料）擷取。包括 LPSM（以及可選地亦包括節目邊界元資料）的各元資料片段係被包括在該位元流之一訊框的無用位元片段中，或者該位元流之一訊框之位元流資訊（“BSI”）片段的“addbsi”欄位，或者在該位元流之

一訊框結束時的輔助資料欄位中（例如，在圖 4 中所示的 AUX 片段）。該位元流之一訊框包括一或兩個元資料片段，其中各個包括 LPSM，且假如該訊框包括兩個元資料片段，一個存在於該訊框的 addbsi 欄位中，且另一個在該訊框的 AUX 欄位中。在一些實施例中，包括 LPSM 的各元資料片段包括具有以下格式的 LPSM 酬載（或容器）片段：

一標頭，（基本上包括一同步字，其識別 LPSM 酬載的開頭，其後接著識別值，例如，LPSM 格式版本、長度、時期、數目與子串流相關值，如在下文表 2 中所指出），以及

在該標頭以後，

至少一個對話指示值（例如，表 2 的參數“對話頻道”），其指示對應的音頻資料是否指示對話或不指示對話（例如，哪些對應音頻資料的頻道指示對話）；

至少一個響度規則遵守值（例如，表 2 的參數“響度規則類型”），其指示出對應的音頻資料是否遵守一指定組響度規則；

至少一個響度處理值（例如，表 2 參數“對話閘控響度校正旗標”、“響度校正類型”的其中一個或多個），其指示至少一種類型的響度處理，其已經在該對應的音頻資料上施行；以及

至少一個響度值（例如，表 2 參數“ITU 相關閘控響度”、“ITU 演講閘控響度”、“ITU (EBU 3341) 短期

3s 響度”與“真實峰值”的其中一個或多個），其指示該對應音頻資料的至少一個響度（例如，峰值或平均響度）特徵。

在一些實施例中，包含 LPSM 與節目邊界元資料的各元資料片段包含一核心標頭（以及可選地亦包含額外核心元件），以及在該核心標頭（或者該核心標頭與其他核心元件）以後，LPSM 酬載（或容器）片段具有以下格式：

一標頭，基本上包括至少一個識別值（例如，LPSM 格式版本、長度、時期、數目與子串流相關值，如在下文表 2 中所指出），以及

在該標頭以後，LPSM 與節目邊界元資料。該節目邊界元資料包括一節目邊界訊框數目，以及一代碼值（例如，一“偏移_存在”值），其指示該訊框是否僅包括一節目邊界訊框數或者一節目邊界訊框數與一偏移值兩者，以及（在一些情形中）一偏移值。

在一些實施程序中，剖析器 205（及/或解碼器台 202）係被組態以從該位元流之一訊框的無用位元片段或“addbsi”欄位或輔助資料欄位擷取各元資料片段，其具有以下格式：

一核心標頭（基本上包括識別該元資料片段之開頭的一同步字，其後接著至少一個識別值，例如，核心元件版本、長度、與時期、延伸元件數目與子串流相關值，其係在下文表 1 中指出）；以及

在該核心標頭以後，至少一個保護值（例如，表 1 的

HMAC 摘要與音頻指紋值），其係對響度處理狀態元資料或者該對應音頻資料之至少一個之解密、鑑定或驗證的至少一個有用；以及

亦在該核心標頭以後，假如該元資料片段包括 LPSM、LPSM 酬載識別（“ID”）與 LPSM 酬載尺寸值，其係將以下的元資料識別為 LPSM 酬載並且指示 LPSM 酬載的尺寸。

LPSM 酬載（或容器）片段（較佳地具有上文具體指定的格式）接著 LPSM 酬載 ID 與 LPSM 酬載尺寸值。

更一般地，由本發明較佳實施例所產生的已編碼音頻位元流具有一結構，該結構提供一機制，以將元資料元件與子元件標示為核心（必備）或延伸（可選元件）。這允許位元流的資料速率（包括它的位元流）在許多應用定標。較佳位元流語法的核心（必備）元件亦應該能夠發出與音頻內容有關的延伸（可選）元件存在（帶內）及/或在遠端位置（帶外）之訊號。

核心元件必須存在於該位元流的各訊框中。核心元件的一些子元件是可選地，並且存在於任何組合中。延伸元件不需要存在於各訊框中（以間接限制位元速率）。因此，延伸元件可存在於一些訊框中，而不是其他。延伸元件的一些子元件是可選的，並且可存在於任何組合中，然而，延伸元件的一些子元件為必備（亦即，假如該延伸元件存在於該位元流的一訊框中）。

在一種類的實施例中，包含一序列音頻資料片段與元

資料片段的已編碼音頻位元流會被產生（例如，藉由實施本發明的音頻處理單元）。該音頻資料片段指示音頻資料，至少一些元資料片段的各個均包括響度處理狀態元資料（LPSM）以及可選地亦包括節目邊界元資料，且該等音頻資料片段係用元資料片段來分時多工。在此種類的較佳實施例中，該等元資料片段的各個具有在本文中所說明的較佳格式。

在一個較佳格式中，已編碼位元流為 AC-3 位元流或 E-AC-3 位元流，且包括 LPSM 的各元資料片段，其係當作額外位元流資訊地（例如藉由編碼器 100 之較佳實施程序的台 107）被包括在該位元流之一訊框之位元流資訊（“BSI”）片段的 “addbsi” 欄位中（圖 6 所示），或者該位元流之一訊框的輔助資料欄位中，或者在該位元流之一訊框的無用位元片段中。

在較佳格式中，在該訊框的 addbsi 欄位（或無用位元片段）中，各訊框均包括具有以下表 1 所示格式的核心元件：

表 1

參數	說明	必備/可選
SYNC [ID]	該同步字為 16 位元值，其值被設定為 0x5838	M
核心元件版本		M
核心元件長度		M
核心元件時期 (xxx)		M
延伸元件數	指示與該核心元件有關之延伸元資料元件的數目。隨著該位元流從生產通過分布與最後排放，此值會增加/減少。	M
子串流結合	說明該核心元件與那(些)子串流有關	M
簽署 (HMAC 摘要)	在整個訊框的音頻資料、核心元件與全部延伸元件上計算的 256-位元 HMAC 摘要(使用 SHA-2 演算法)。	M
PGM 邊界倒數	欄位僅僅出現於在音頻節目檔案/流之頭部或尾部的一些數目訊框。因此，核心元件版本改變可被使用來發出包括此參數的訊號。	O
音頻指紋	音頻指紋接管該核心元件期間欄位所代表的一些數目 PCM 音頻樣本。	O
視頻指紋	視頻指紋接管一些數目的壓縮視頻樣本(假如有的話)，其係由核心元件時期欄位所代表。	O
URL (全球資源定位器) / UUID (通用唯一識別碼)	此欄位係被定義以攜帶 URL 及/或 UUID(它對指紋是多餘的)，其係參考額外節目內容(要素)的外部位置及/或與該位元流相關的元資料。	O

在較佳格式中，包含 LPSM 之 addbsi（或輔助資料）欄位或無用位元片段的各個均包含核心標頭（以及可選地亦包含額外核心元件），而且在該核心標頭（或該核心標頭與其他核心元件）以後，以下的 LPSM 值（參數）：

酬載 ID（識別該元資料為 LPSM），接著該核心元件值（例如，如表 1 所具體指示）：

酬載尺寸（指示 LPSM 酬載的尺寸），接著該酬載 ID；以及

LPSM 資料（接著該酬載 ID 與酬載尺寸值），其具有在以下表（表 2）中所指示的格式：

表 2

LPSM 參數 [智慧型響度]	說明	單一狀態 的數目	必備/可選	插入速率 (參數的更新週期)
LPSM 版本			M	
LPSM 時期 (xxx)	僅可應用到 xxx 欄位		M	
LPSM 數			M	
LPSM 子流結合			M	
對話頻道	指示 L、C&R 音頻頻道的哪種組合包含在先前 0.5 秒的演講。當演講不存在於任何 L、C 或 R 組合時，那麼此參數將指示“沒有對話”	8	M	-0.5 秒 (基本上)
響度規則類型	指示該相關的音頻資料流遵守一具體組規則(例如，ATSCA/85 或 EBU128)	8	M	訊框

對話閘控響度校正旗標	指示假如該相關音頻流已經以對話閘控為基礎被校正	2	O(只有在假如響度_規則_類型指示出該對應音頻未被校正之下存在)	訊框
響度校正類型	指示是否該相關音頻流已經用有限預見(以檔案為基礎)或用即時(RT)響度與動態範圍控制器被校正	2	O(只有在假如響度_規則_類型指示出該對應音頻未被校正之下存在)	訊框
ITU 相關閘控響度 (INF)	指示所施加之該相關音頻流 w/o 元資料的 ITU-R BS.1770-3 整合響度(例如, 7 位元: -58->+5.5 LKFS 0.5 LKFS 步驟)	128	O	1 秒
ITU 演講閘控響度 (INF)	指示所施加之該相關音頻流 w/o 元資料之該演講/對話的 ITU-R BS.1770-1/3 整合響度(例如, 7 位元: -58->+5.5 LKFS 0.5 LKFS 步驟)	128	O	1 秒
ITU (EBU 3341) 短期 3s 響度	指示所施加之該相關音頻流 w/o 元資料(滑動窗口)之 3 秒未閘控 ITU (ITU-BS.1771-1) 響度@~10Hz 插入速率(例如, 8 位元: 116->+11.5 LKFS 0.5 LKFS 步驟)	256	O	0.1 秒

真實峰值的值	指示所施加之該相關音頻流 w/o 元資料之 ITU-RBS.1770-3 Annex2 真實峰值的值(dBTP) (亦即, 在元件時期欄位中所發出訊號之訊框時期上的最大值) 116->+ 11.5 LKFS 0.5 LKFS 步驟	256	0	0.5 秒
下游偏移	指示下游響度偏移			
節目邊界	在訊框中指示何時節目邊界將或已經發生。當節目邊界不在訊框邊界上時, 可選的樣本偏移將指示在訊框真實節目邊界中發生多遠。			

在根據本發明所產生之已編碼位元流的另一較佳格式中, 該位元流為 AC-3 位元流或 E-AC-3 位元流, 且包括 LPSM (以及可選地亦包括節目邊界元資料) 的各元資料片段係 (例如, 藉由編碼器 100 之較佳實施程序的台 107) 被包括在以下任一者: 該位元流之一訊框的無用位元片段; 或者該位元流之一訊框之位元流資訊 ("BSI") 片段的 "addbsi" 欄位 (如圖 6 所示); 或者在該位元流之一訊框結束時的一輔助資料欄位 (例如, 圖 4 所示的 AUX 片段)。一訊框可包括一或兩個元資料片段, 其中每個均包括 LPSM, 且假如該訊框包括兩個元資料片段, 一個可存在於該訊框的額外位元流 (addbsi)

欄位中，且另一個在該訊框的 AUX 欄位中。包括 LPSM 的各元資料片段具有參考上文表 1 與 2 而在上文具體指定的格式（亦即，它包括在表 1 所具體指定的核心元件，其後接著酬載 ID（識別該元資料為 LPSM）以及上文具體指定的酬載尺寸值，其後接著酬載（LPSM 資料，其具有在表 2 中所指示的格式）。

在另一個較佳格式中，已編碼位元流為杜比 E 位元流，且包括 LPSM（以及可選地亦包括節目邊界元資料）的各元資料片段為杜比 E 保護帶間隔的前 N 個樣本位置。包括此類包括 LPSM 之元資料片段的杜比 E 位元流較佳地包括一值，其指示以 SMPTE 337M 前文之 Pd 字來發訊號的 LPSM 酬載長度（SMPTE 337M Pa 字重複率較佳地維持與相關音頻訊框率相等）。

在較佳格式中，其中已編碼位元流為 E-AC-3 位元流，包括 LPSM（以及可選地亦包括節目邊界元資料）之各元資料片段，其係當作額外位元流資訊地（例如，藉由編碼器 100 之較佳實施程序的台 107）被包括在該位元流之一訊框的無用位元片段中或者該位元流資訊（“BSI”）片段的“addbsi（額外位元流）”欄位中。我們接著說明在此較佳格式中、用 LPSM 來編碼 E-AC-3 位元流的額外態樣：

1. 在 E-AC-3 位元流產生期間內，雖然 E-AC-3 編碼器（其將該等 LPSM 值插入到位元流內）為“活性”，就所產生的各訊框（同步訊框）而言，該位元流應該包括在該

訊框之 `adbsi` (額外位元流) 欄位 (或者無用位元片段) 中所攜帶的元資料方塊 (包括 `LPSM`)。攜帶元資料方塊所需要的位元不應該增加編碼器位元速率 (訊框長度)；

2. 各元資料方塊 (包含 `LPSM`) 應該包含以下資訊：

`響度_校正_類型_旗標`：在此 '1' 指示對應音頻資料的響度在從該編碼器的上游被校正，且 '0' 指示響度係由被嵌入編碼器中的響度校正器所校正 (例如，圖 2 之編碼器 100 的響度處理器 103)；

`演講_頻道`：指示哪 (些) 來源頻道包含演講 (在先前 0.5 秒內)。假如沒有檢測出任何演講的話，此值將指示其本身；

`演講_響度`：指示包含演講之各對應音頻頻道的整合演講響度 (在先前 0.5 秒內)；

`ITU_響度`：指示各對應音頻頻道的整合 ITU BS.1170-3 響度；以及

`增益`：用於解碼器中之反轉的響度合成增益 (以顯示可逆性)；

3. 雖然 `E-AC-3` 編碼器 (其係將 `LPSM` 值插入到位元流內) 為 "活性" 並且以 '信賴' 旗標來接收 `AC-3` 訊框，在編碼器中的響度控制器 (例如，圖 2 之編碼器 100 的響度處理器 103) 應該被旁路。該 '信賴' 來源對話歸一化與 `DRC` 值應該 (例如藉由編碼器 100 的產生器 106) 直通到 `E-AC-3` 編碼器組件 (例如，編碼器 100 的台 107)。 `LPSM` 方塊產生會持續且 `響度_校正_類型_旗標` 被設定在

‘1’。響度控制器旁路序列必須被同步化到‘信賴’旗標出現之解碼 AC-3 訊框的開頭。響度控制器旁路序列應該實施如下：在 10 個音頻方塊時期內（亦即，53.3 毫秒），校平器_數量控制會從 9 的值減少到 0 的值，且校平器_往回_結束_計控制係被放置在旁路模式內（此操作應該導致無縫轉換）。該校平器之術語‘信賴’旁路意味著該來源位元流的對話歸一化值亦可在該編碼器的輸出上被重新利用（例如，假如該‘信賴’的來源位元流具有-30 的對話歸一化值，那麼編碼器的輸出則應該使用-30 為輸出的對話歸一化值）；

4. 雖然 E-AC-3 編碼器（其係將 LPSM 值插入到位元流內）為“活性”並且接收不具有‘信賴’旗標的 AC-3 訊框，嵌入在編碼器中的響度控制器（例如，圖 2 之編碼器 100 的響度處理器 103）應該是活性的。LPSM 方塊產生會持續且響度_校正_類型_旗標被設定在‘0’。響度控制器活化序列必須被同步化到‘信賴’旗標消失之解碼 AC-3 訊框的開頭。響度控制器活化序列應該實施如下：在 1 個音頻方塊時期內（亦即，5.3 毫秒），校平器_數量控制會從 0 的值增加到 9 的值，且校平器_往回_結束_計控制係被放置在‘活性’模式內（此操作應該導致無縫轉換並且包括往回_結束_計整合重新設定）；以及

5. 在編碼期間內，圖形使用者介面（GUI）應該將以下參數指示給使用者：“輸入音頻節目：[信賴/不信賴]” — 此參數的狀態係以輸入訊號內之“信賴”旗標的

存在為基礎；以及“即時響度校正：[致能/失能]”此一參數的狀態係以是否嵌入在編碼器中的此響度控制器是活性為基礎。

當將被包括在該位元流之各訊框之無用位元片段中或位元流資訊（“BSI”）片段的“額外位元流”欄位中之具有（呈較佳格式）LPSM的AC-3或E-AC-3位元流解碼時，該解碼器應該剖析（該無用位元片段或額外位元流欄位中的）LPSM方塊資料並且將全部該擷取LPSM值直通到圖形使用者介面（GUI）。該組所擷取的LPSM值係每逢一訊框地更新。

在根據本發明所產生之已編碼位元流的另一較佳格式中，已編碼位元流為AC-3位元流或E-AC-3位元流，且包括LPSM的各元資料片段，其係（例如藉由編碼器100之較佳實施程序的台107）被包括在該位元流之一訊框之一無用位元片段中、或一Aux片段中或當作額外位元流資訊地被包括在位元流資訊（“BSI”）片段的“addbsi（額外位元流）”欄位中（圖6所示）。在本格式中（其為參考表1與2而在上文所說明之格式上的變化），包含LPSM之addbsi（或Aux或無用位元）欄位的各個包含以下LPSM值：

在表1中所具體指定的核心元件，其後接著酬載ID（識別該元資料為LPSM）與酬載尺寸值，其後接著酬載（LPSM資料），其具有以下格式（類似在上文表2中所指示的必備元件）：

LPSM 酬載的版本：一個 2 位元欄位，其指示 LPSM 酬載之版本；

dialchan：一個 3 位元欄位，其指示對應音頻資料之左、右及/或中央頻道是否包含口說對話。該 **dialchan** 欄位的位元配置如下：位元 0，其指示在左頻道對話的存在，其係儲存在 **dialchan** 欄位的最高有效位元中；以及位元 2，其指示出對話存在於中央頻道中，其係儲存在 **dialchan** 欄位之最低有效位元中。假如在該節目的先前 0.5 秒期間內，該對應頻道包含口說對話的話，該 **dialchan** 欄位的各位元則被設定於‘1’。

loudregtyp：一個 4 位元欄位，其指示該節目響度遵守哪種響度規則標準。將“**loudregtyp**”欄位設定在“000”，指示 LPSM 沒有指示響度規則順從性。例如，此欄位的一個值（例如，0000）可指示順從響度規則標準沒被指示，本欄位的另一個值（例如，0001）可指示該節目的音頻資料遵守 ATSC A/85 標準，且本欄位的另一個值（例如，0010）指示該節目的音頻資料遵守 EBU R128 標準。在該實例中，假如該欄位被設定在除了‘0000’以外的任何值，**loudcorrdialgat** 與 **loudcorrtyp** 欄位應該在該酬載中接在其後：

loudcorrdialgat：一位元欄位，其指示對話閘控響度校正是否已經被施加。假如該節目的響度已經使用對話閘控被校正，**loudcorrdialgat** 欄位的值則被設定在‘1’。不然，它會被設定在‘0’；

loudcorrtyp：一位元欄位，其指示被施加到該節目之響度校正的類型。假如該節目的響度已經用無限預見（以檔案為基礎）的響度校正過程來校正，**loudcorrtyp** 欄位的值則可設定在‘0’。假如該節目的響度已經使用即時響度測量與動態範圍控制之組合來校正，此欄位的值則被設定在‘1’；

loudrelgate：一位元欄位，其指示相關閘控響度資料（ITU）是否存在。假如 **loudrelgate** 欄位被設定在‘1’，7 位元的 **ituloudrelgat** 欄位應該在該酬載中接在其後；

loudrelgat：7-位元欄位，其指示相關閘控節目響度（ITU）。此欄位指示該音頻節目的整合響度，其係在沒有起因於被施加之對話歸一化與動態範圍壓縮的任何增益調整之下根據 ITU-R BS.1770-3 來測量。0 至 127 的值係被詮釋為在 0.5LKFS 步驟中的 -58LKFS 至 +5.5LKFS。

loudspchgate：一位元欄位，其指示演講閘控響度資料（ITU）是否存在。假如 **loudspchgate** 欄位被設定在‘1’，7 位元 **loudspchgat** 欄位應該在該承載中接在其後；

loudspchgat：7 位元欄位，其指示演講閘控節目響度。此欄位指示該整個對應音頻節目的整合響度，其係根據 ITU-R BS.1770-3 的公式（2）來測量而且沒有起因於被施加之對話歸一化與動態範圍壓縮的任何增益調整。在 0.5LKFS 步驟中，0 至 127 的值係被詮釋為 -58 至 +5.5LKFS。

loudstrm3se：一位元欄位，其指示短期（3 秒）響度

資料是否存在。假如該欄位被設定在‘1’，7 位元 `loudstrm3s` 欄位應該在該承載中接在其後；

`loudstrm3s`：7 位元欄位，其指示對應音頻節目之先前 3 秒的未閘控響度，其係根據 ITU-R BS.1771-1 來測量而且沒有起因於被施加之對話歸一化與動態範圍壓縮的任何增益調整。在 0.5LKFS 步驟中，0 至 256 的值係被詮釋為 -116LKFS 至 +11.5LKFS；

`truepke`：一位元欄位，其指示真實峰值響度資料是否存在。假如 `truepke` 欄位被設定於‘1’，8 位元 `truepke` 欄位應該在該承載中接在其後；以及

`truepk`：8 位元欄位，其指示該節目的真實峰值樣本值，其係根據 ITU-R BS.1770-3 的 Annex 2 來測量而且沒有起因於被施加之對話歸一化與動態範圍壓縮的任何增益調整。在 0.5LKFS 步驟中，0 至 256 的值係被詮釋為 -116 LKFS 至 +11.5LKFS。

在一些實施例中，在 AC-3 位元流或 E-AC-3 位元流之訊框之無用位元片段或輔助資料（或“`addbsi`”）欄位之元資料片段的核​​心元件包含一核​​心標頭（基本上包括識別值，例如核​​心元件版本），而且在該核​​心標頭以後：指示出指紋資料（或者其他保護值）是否被包括以用於該元資料片段之元資料的值、指示出是否外部資料（與對應該元資料片段之元資料的音頻資料相關）存在的值、由該核​​心元件所識別之各類型元資料（例如，LPSM、及/或除了 LPSM 以外之一類型的元資料）的酬載 ID 與酬載尺寸

值、與由該核心元件所識別之至少一類型元資料的保護值。該元資料片段的元資料酬載接著該核心標頭，並且（在一些情形中）被套入該核心元件的值內。

本發明的典型實施例包括在有效方式中之已編碼音頻位元流中的節目邊界元資料，其允許在藉由該位元流所指示之連續音頻節目之間之至少一個邊界的準確與穩定決定。典型的實施例在它們允許準確節目邊界決定的意義中允許一節目邊界準確與穩定的決定，甚至假使其中指示不同節目的位元流以截斷該等接合位元流其中一個或兩個的方式被接合在一起（以產生本發明位元流）（並且因此放棄已經被包括在至少一個事先接合位元流中的節目邊界元資料）。

在典型的實施例中，在本發明位元流訊框中的節目邊界元資料為指示一訊框數目的一節目邊界旗標。基本上，該旗標指示在目前訊框（包括該旗標的該訊框）與節目邊界（該目前音頻節目的開始或結束）之間的訊框數目。在一些較佳實施例中，節目邊界旗標以對稱、有效的方式在指示單一節目之各位元流片段開始與結束時被插入（亦即，在該片段開始以後之發生於一些預定數目訊框內的訊框中，以及在該片段結束以前之發生於一些預定數目訊框內的訊框中），以致於當兩個此位元流片段被連接時（以便指示一序列的兩個節目），該節目邊界元資料可（例如，對稱地）存在於該兩節目之間之邊界的兩側上。

最大穩定性可藉由插入節目邊界旗標於指示一節目之

位元流的各訊框中來得到，但是由於資料速率的相關增加，這基本上不實用。在典型的實施例中，節目邊界旗標係被插入於該已編碼音頻位元流的僅一子集訊框中（其指示一個音頻節目或者一序列音頻節目），而且該邊界旗標插入速率為各位元流訊框（其中插入一旗標）距最近該各訊框之該節目邊界之增加間距的非增加函數，在此“邊界旗標插入速率”指示包括節目邊界旗標之訊框數目（指示出一節目）對不包括節目邊界旗標之訊框數目（指示出該節目）的平均比率，在此，該平均值為已編碼音頻位元流之若干（例如，相當小數目）連續訊框上的移動平均值。

增加該邊界旗標插入速率（例如，在較靠近節目邊界之位元流中的位置）會增加傳送該位元流所必要的資料速率。為了補償這，各插入旗標的尺寸（位元數目）會隨著邊界旗標插入速率的增加而較佳地減少（例如，以致於在該位元流之第“N”個訊框中（在此，N是整數）節目邊界旗標的尺寸，其為在第“N”個訊框與最近節目邊界之間距離（訊框數目）的非增加函數）。在一種類的實施例中，邊界旗標插入速率為距該最近節目邊界之（各旗標插入位置的）增加距離的對數減少函數，且就包括其中一個旗標的各包含旗標訊框而言，在該包含旗標訊框中之旗標的尺寸，其等於或大於位置比該包含旗標訊框更靠近最近節目邊界之一訊框中的各旗標的尺寸。基本上，各旗標尺寸係藉由從該旗標插入位置到最近節目邊界之訊框數目的增加函數所決定。

例如，考慮圖 8 與 9 的實施例，其中由一訊框數（在最頂列）所識別的各行指示出已編碼音頻位元流的一訊框。該位元流指示出具有第一節目邊界（指示出該節目開始）的音頻節目，其立即發生在由圖 9 左側上之訊框數“17”所識別之該行的左邊，以及具有第二節目邊界（指示出該節目的結束），其立即發生在由在圖 8 右側上之訊框數“1”所識別之該行的右邊。被包括在圖 8 所示訊框中的節目邊界旗標倒數在目前訊框與第二節目邊界之間的訊框數目。被包括在圖 9 所示訊框中的節目邊界旗標計數在目前訊框與第一節目邊界之間的訊框數目。

在圖 8 與 9 的實施例中，節目邊界旗標係在該位元流所指示之音頻節目開頭以後僅僅插入於該已編碼位元流之前 X 個訊框之第“ 2^N ”個訊框的各個中，以及在最靠近由該位元流所指示之節目結束時（該位元流之最後 X 個訊框的）之第“ 2^N ”個訊框的各個中，在此該節目包含 Y 個訊框， X 為小於或等於 $Y/2$ 的整數，且 N 為從 1 至 $\log_2(X)$ 之範圍中的正整數。因此（如圖 8 與 9 中所指示），節目邊界旗標被插入於位元流的第二訊框中（ $N=1$ ）（最靠近該節目開始的該包含旗標訊框），第四訊框中（ $N=2$ ），第八訊框中（ $N=3$ ）等等，以及自該位元流結束的第八訊框中、自該位元流結束的第四訊框中、以及自該位元流結束時的第二訊框中（最靠近該節目結束時的該包含旗標訊框）。在本實例中，在自該節目開始（或結束）之第“ 2^N ”個訊框中的節目邊界旗標包含 $\log_2(2^{N+2})$ 二元位

元，如圖 8 與 9 所指示。因此，在自該節目開始（或結束）之第二個訊框（ $N=1$ ）中的節目邊界旗標包含 $\log_2(2^{N+2})=\log_2(2^3)=3$ 個二元位元，且在自該節目開始（或結束）之第四個訊框（ $N=2$ ）中的旗標包含 $\log_2(2^{N+2})=\log_2(2^4)=4$ 個二元位元等等。

在圖 8 與 9 的實例中，各節目邊界旗標的格式係如下。各節目邊界旗標係由一引導“1”位元、在該引導位元以後的一序列“0”位元（沒有任何“0”位元或一或多個連續“0”位元）與兩位元尾碼所組成。如圖 8 所示，就該位元流之最後 X 訊框中（最靠近節目結束的該等訊框）的旗標而言，尾碼為“11”。如圖 9 所示，就該位元流之前 X 訊框中（最靠近節目開始的該等訊框）的旗標而言，尾碼為“10”。因此，為了讀取（解碼）各旗標，在引導“1”位元與尾碼之間的零的數目會被計算。假如該尾碼被識別為“11”，該旗標則指示出在目前訊框（包括該旗標的該訊框）與該節目結束之間有 $(2^{Z+1}-1)$ 個訊框，在此 Z 為在該旗標的引導“1”位元與尾碼之間零的數目。該解碼器可被有效地實施，以忽略各此旗標的第一與最後位元，以決定該旗標之其他（中間）位元序列的反轉（例如，假如該中間位元序列為“0001”，“1”位元為在該序列中的最後位元，該中間位元反轉序列為“1000”，“1”位元為在該反轉序列中的第一位元），而且將該中間位元之反轉序列的二元值識別為目前訊框（包括該旗標的訊框）相關於節目結束的指數。例如，假

如該中間位元的反轉序列為“1000”，此反轉序列會具有二元值 $2^4=16$ ，且該訊框會被識別為在該節目結束以前的第 16 個訊框（如在圖 8 的該行中所指示，其說明訊框“0”）。

假如該尾碼被識別為“10”，該旗標則指示在該節目開頭與目前訊框（包括該旗標的該訊框）之間有 $(2^{Z+1}-1)$ 個訊框，在此 Z 為在該旗標的引導“1”位元與尾碼之間的零的數目。該解碼器可被有效地實施，以忽略各此旗標的第一與最後位元，以決定該旗標之中間位元之序列的反轉（例如，假如該中間位元序列為“0001”，“1”位元為在該序列中的最後位元，該中間位元反轉序列為“1000”，“1”位元為在該反轉序列中的第一位元），而且將該中間位元反轉序列的二元值識別當作目前訊框（包括該旗標的訊框）相關於節目開始的指數。例如，假如該中間位元反轉序列為“1000”，此反轉序列具有二元值 $2^4=16$ ，且該訊框會被識別當作在該節目開始以後的第 16 個訊框（如在圖 9 的該行中所指示，其說明訊框“32”）。

在圖 8 與 9 的實例中，在位元流所指示之音頻節目開頭以後，節目邊界旗標係僅僅存在於一已編碼位元流之前 X 個訊框之第“ 2^N ”個訊框的各個中，並且在最靠近由該位元流所指示之節目結束時（該位元流之最後 X 個訊框的）之第“ 2^N ”個訊框的各個中，在此該節目包含 Y 個訊框， X 為小於或等於 $Y/2$ 的整數，且 N 為從 1 至 \log_2

(X) 之範圍中的正整數。包括該節目邊界旗標僅僅添加 1.875 位元/訊框的平均位元速率到傳送沒有旗標之該位元流所必要的位元速率。

在圖 8 與 9 之實施例的典型實施程序中，其中該位元流為 AC-3 已編碼音頻位元流，各訊框包含 1536 個數位音頻樣本的音頻內容與元資料。就 48kHz 的取樣速率而言，這代表 32 毫秒的數位音頻或者每秒音頻 31.25 訊框的速率。因此，在此類實施例中，在與一節目邊界隔開一些數目訊框（“X” 個訊框）之一訊框中的節目邊界旗標，其指示在該包含旗標訊框結束以後，該邊界發生 32X 毫秒（或者在包含旗標訊框開頭以前，32X 毫秒）。

在圖 8 與 9 之實施例的典型實施程序中，其中該位元流為 E-AC-3 已編碼音頻位元流，該位元流的各訊框包含 256、512、768 或 1536 個數位音頻樣本的音頻內容與元資料，其係依據該訊框是否各自包含一個、兩個、三個或六個方塊的音頻資料。就 48kHz 的取樣速率而言，這各自代表 5.333、10.667、16 或 32 毫秒的數位音頻或者每秒音頻各自 189.9、93.75、62.5 或 31.25 訊框的速率。因此，在此類實施例中（假定各訊框指示 32 毫秒的數位音頻），在與一節目邊界分開一些數目訊框（“X” 個訊框）之一訊框中的節目邊界旗標，其指示在該包含旗標訊框結束以後，該邊界發生 32X 毫秒（或者在包含旗標訊框開頭以前，32X 毫秒）。

在一些實施例中，其中一節目邊界發生在一音頻位元

流的訊框內（亦即，沒有對準一訊框的開始或結束），被包括在該位元流之一訊框的節目邊界元資料包括節目邊界訊框數（亦即，指示出在該包含-訊框數之訊框的開始或結束與節目邊界之間的全訊框數目的元資料）與偏移值。該偏移值指示在包含-節目邊界訊框的開始或結束與在該包含-節目邊界訊框內之節目邊界的真實位置之間的偏移（基本上，為許多個樣本）。

已編碼音頻位元流指示出一對應序列音頻節目的一序列節目（聲道），而且此音頻節目的邊界傾向於發生在視頻訊框的邊緣而非在音頻訊框的邊緣上。同樣地，一些音頻編解碼器（例如，E-AC-3 編解碼器）使用沒有對準視頻訊框的音頻訊框尺寸。同樣地，在一些情形中，最初編碼的音頻位元流經歷轉碼，以產生轉碼位元流，且最初編碼的位元流具有與該轉碼位元流不同的訊框尺寸，以致於節目邊界（由最初已編碼位元流所決定）不會被保證發生在該轉碼位元流的訊框邊界上。例如，假如最初已編碼位元流（例如，圖 10 的位元流“IEB”）具有每一訊框 1536 個樣本的訊框尺寸，且該轉碼位元流（例如，圖 10 的位元流“TB”）具有每一訊框 1024 個樣本的訊框尺寸，該轉碼過程則會造成真實的節目邊界沒有發生在該轉碼位元流的訊框邊界上，但卻在其一訊框的某處（例如，如圖 10 所指示，512 個樣本到該轉碼位元流的一訊框內），其係由於不同編解碼器的不同訊框尺寸。本發明的實施例，其中被包括在已編碼音頻位元流之訊框中的節目

邊界元資料包括一偏移值以及一節目邊界訊框數，其係對在本段落中所記錄的三種情形（以及其他情形）有用。

參考圖 8 與 9 而在以上說明的實施例不包括一偏移值（例如，一偏移欄位）於該已編碼位元流的任一訊框中。在本實施例上的變化中，一偏移值係被包括在包括節目邊界旗標之已編碼音頻位元流的各訊框中（例如，在對應圖 8 之編號 0、8、12 與 14 之訊框與圖 9 之編號 18、20、24 與 32 之訊框的訊框中）。

在一種類的實施例中，一資料結構（在包含本發明節目邊界元資料之已編碼位元流的各訊框中）包括一代碼值，其指示該訊框是否僅僅包括一節目邊界訊框數或者一節目邊界訊框數與一偏移值兩者。例如，該代碼值為單位元欄位（在本文中稱為“偏移_存在”欄位）的值，該值“偏移_存在”=0 指示出沒有任何偏移值被包括在該訊框中，且該值“偏移_存在”=1 指示出節目邊界訊框數與偏移值兩者係被包括在該訊框中。

在一些實施例中，AC-3 或 E-AC-3 已編碼音頻位元流的至少一個訊框包括一元資料片段，該元資料片段包括 LPSM 與節目邊界元資料（以及可選地亦包括其他元資料），以用於由該位元流所決定的音頻節目。各此類的元資料片段（其係被包括在該位元流的 `adbsi` 欄位、或輔助資料欄位或無用位元片段中）包含一核心標頭（以及可選地亦包括額外核心元件），以及在該核心標頭（或者該核心標頭與其他核心元件）以後，LPSM 酬載（或容器）

片段具有以下格式：

一標頭，（基本上包括至少一個識別值，例如，LPSM 格式版本、長度、時期、數目與子串流相關值），以及

在該標頭以後，節目邊界元資料（其係可包括一節目邊界訊框數目、一代碼值（例如，一“偏移_存在”值），其指示出該訊框是否僅僅包括一節目邊界訊框數或者一節目邊界訊框數與一偏移值兩者，以及在一些情形中，一偏移值）與 LPSM。LPSM 包括：

至少一個對話指示值，其指示對應的音頻資料是否指示對話或不指示對話（例如，對應音頻資料的哪些頻道指示對話）。該對話指示值可指示對話是否存在於該對應音頻資料之頻道的任一組合或全部；

至少一個響度規則遵守值，其指示對應的音頻資料是否遵守一指定組響度規則；

至少一個響度處理值，其指示至少一種響度處理，其已經在該對應的音頻資料上施行；以及

至少一個響度值，其指示該對應音頻資料的至少一個響度（例如，峰值或平均響度）特徵。

在一些實施例中，LPSM 酬載片段包括指示該訊框是否僅僅包括一節目邊界訊框數或者一節目邊界訊框數與一偏移值兩者的一代碼值（一“偏移_存在”值）。例如，在一項此類實施例中，當此類代碼值指示出（例如，當偏移_存在=1 時）該訊框包括一節目邊界訊框數與一偏移值

時，LPSM 酬載片段包括一偏移值，其為 11 位元的未簽署整數（亦即，具有從 0 至 2048 的值），且其指示在已發訊號訊框邊界（包括節目邊界之訊框的邊界）與真實節目邊界之間的額外音頻樣本的數目。假如該節目邊界訊框數指示出到包含-節目邊界訊框的訊框數目（以現在的訊框速率），節目邊界的精確位置（以樣本數目為單位）（相關於包括 LPSM 酬載片段之訊框的開頭或結束）應該計算如下：

$$S = (\text{訊框}_\text{計數器} * \text{訊框尺寸}) + \text{偏移},$$

在此，S 為到節目邊界的樣本數目（從包括 LPSM 酬載片段之訊框的開頭或結束），“訊框_計數器”為由節目邊界訊框數所指示的訊框數，“訊框尺寸”為每一訊框的樣本數目，且“偏移”為由該偏移值所指示出的樣本數目。

其中節目邊界旗標之插入速率在靠近真實節目邊界之處增加的一些實施例實施一規則，該規則為假如該訊框小於或等於來自包括該節目邊界之訊框的一些數目（“Y”）的訊框，一偏移值從不被包括在訊框中。基本上，Y=32。就實施此規則（Y=32）的 E-AC-3 編碼器而言，編碼器從不插入一偏移值於音頻節目的最後秒中。在本情形中，該接收裝置負責維持計時器，並且因此施行它本身的偏移計算（回應在超過來自包含節目邊界訊框之 Y 訊框的已編碼位元流之訊框中的節目邊界元資料，包括偏移值）。

就其音頻節目被視為“訊框對準”對應視頻節目之視頻訊框的節目而言（例如，以杜比 E 已編碼音頻的典型貢獻饋給），將偏移值包括於指示音頻節目的已編碼位元流中將是多餘的。因此，偏移值基本上將不被包括在此類已編碼位元流中。

參考圖 11，我們接著考慮其中已編碼音頻位元流被接合在一起以產生本發明音頻位元流實施例的情形。

在圖 11 頂部的位元流（標為“情境 1”）指示出包括節目邊界元資料（節目邊界旗標，F）的整個第一音頻節目（P1），其後接著亦包括節目邊界元資料（節目邊界旗標，F）的整個第二音頻節目（P2）。在第一節目之結束部分的節目邊界旗標（其中有些係被顯示於圖 11）等於或類似那些參考圖 8 所說明者，其係並且決定在該兩節目之間的邊界位置（亦即，在第二節目開始的邊界）。在第二節目開始部分中的節目邊界旗標（其中一些顯示於圖 11）等於或類似那些參考圖 9 來說明者，而且他們亦決定該邊界的位置。在典型的實施例中，編碼器或解碼器實施計時器（藉由在第一節目中的旗標所校準），其倒數到該節目邊界，且相同計數器（藉由在第二節目中的旗標所校準）則從相同節目邊界計數。如由圖 11 之情境 1 中的邊界計數器圖所指示，此類計數器的倒數（由第一節目中的旗標所校準）會在邊界達到零，而且計數器的計數（由第二節目中的旗標所校準）指示該邊界的相同位置。

來自圖 11 頂部的第二位元流（標為“情境 2”）指

示出包括節目邊界元資料（節目邊界旗標，F）的整個第一音頻節目（P1），其後接著不包括節目邊界元資料的整個第二音頻節目（P2）。在第一節目之結束部分的節目邊界旗標（其中有些係被顯示於圖 11）等於或類似那些參考圖 8 所說明者，其係並且決定在該兩節目之間的邊界位置（亦即，在第二節目開始的邊界），正如在情境 1。在典型的實施例中，編碼器或解碼器實施計時器（藉由在第一節目中的旗標所校準），其倒數到該節目邊界，且相同計時器（沒有被進一步校準）則持續從節目邊界計數（如由圖 11 之情境 2 中的邊界計時器圖所指示）。

來自圖 11 頂部的第三位元流（標為“情境 3”）指示出一截斷的第一音頻節目（P1），其包括節目邊界元資料（節目邊界旗標，F），而且其已經與亦包括節目邊界元資料（節目邊界旗標，F）的整個第二音頻節目（P2）接合。該接合已經移除第一節目的最後“N”個訊框。在第二節目之開始部分的節目邊界旗標（其中有些係被顯示於圖 11）等於或類似那些參考圖 9 來說明者，而且它們決定在該截斷的第一節目與整個第二節目之間的邊界（接合）位置。在典型的實施例中，編碼器或解碼器實施計時器（藉由在第一節目中的旗標所校準），其倒數到該未截斷第一節目結束，且相同計時器（由第二節目中的旗標所校準）則從第二節目開始計數。第二節目的開始係為在情境 3 中的節目邊界。如在圖 11 之情境 3 的邊界計時器圖所指示，此類計時器的倒數（由在第一節目中的節目邊界

元資料所校準) 會在它以經達到零以前 (回應在第一節目中的節目邊界元資料) 被重設 (回應在第二節目中的節目邊界元資料)。因此, 雖然 (藉由剪接) 第一節目之截斷可避免計時器識別在該截斷第一節目與第二節目開始之間的節目邊界, 以單獨回應 (亦即, 受到其校準) 在第一節目中的節目邊界元資料, 在第二節目中的節目元資料會重新設定該計時器, 以致於該重新設定計時器正確地指示在該截斷第一節目與第二節目開始之間之節目邊界的位置 (以當作對應該重新設定計時器之“零”數的位置)。

第四位元流 (標為“情境 4”) 指示出一截斷的第一音頻節目 (P1), 其包括節目邊界元資料 (節目邊界旗標, F), 以及一截斷的第二音頻節目 (P2), 其包括節目邊界元資料 (節目邊界旗標, F) 以及與一部份 (非截斷部分) 的第一音頻節目接合。在整個 (事先截斷) 第二節目之開始部分中的節目邊界旗標 (其中一些顯示於圖 11) 相等或類似那些參考圖 9 來說明者, 而且在整個 (事先截斷) 第一節目 (其中一些顯示於圖 11) 之結束部分中的節目邊界旗標相等或類似那些參考圖 8 來說明者。該接合已經移除第一節目的最後“N”個訊框 (以及因此在剪接以前已經被包括在其中之節目邊界旗標的其中一些) 以及第二節目的前“M”個訊框 (以及因此在剪接以前已經被包括在其中之節目邊界旗標的其中一些)。在典型的實施例中, 編碼器或解碼器實施計時器 (藉由在該截斷的第一節目中的旗標所校準), 其朝該未截斷第一節目結束

來倒數，且相同計時器（由該截斷第二節目中的旗標所校準）則從該未截斷第二節目的開始計數。如在圖 11 之境 4 的邊界計時器圖所指示，此類計時器的倒數（由在第一節目中的節目邊界元資料所校準）會在它以經達到零以前（回應在第一節目中的節目邊界元資料）被重設（回應在第二節目中的節目邊界元資料）。（藉由剪接）第一節目的截斷可避免計時器識別在該截斷的第一節目與該截斷的第二節目開始之間的節目邊界，以單獨回應（亦即，受到其校準）在第一節目中的節目邊界元資料。不過，該重新設定的計時器不會正確地指示在該截斷的第一節目結束與該截斷的第二節目開始之間之節目邊界的位置。因此，兩接合位元流的截斷可避免在它們之間邊界的正確決定。

本發明實施例可用硬體、韌體或軟體或兩者的組合來實施（例如，以當作可程式化邏輯陣列）。除非另外被具體指示，被包括當作本發明之一部分的演算法或過程本質上不會與任何特定電腦或其他設備有關。特別地，各種一般目的之機械可用根據本文中之學說來撰寫的程式來使用，或者它可更方便地架構更專門的設備（例如，積體電路）以施行必要的方法步驟。因此，本發明可在一或多個可程式化電腦系統上執行的一或多個電腦程式中被實施（例如，圖 1 之任一元件或者圖 2 之編碼器 100（或其元件）或圖 3 之解碼器 200（或其元件）或圖 3 之後處理器 300 的實施），各個均包含至少一個處理器、至少一個資料儲存系統（包括揮發性與非揮發性記憶體與/或儲存元

件)、至少一個輸入裝置或埠、以及至少一個輸出裝置或埠。程式代碼係被施加到輸入資料以施行在本文中所說明的功能並且產生輸出資訊。該輸出資訊係以已知的方式被施加到一或多個輸出裝置。

各此類節目係以任何希望的電腦語言來實施(包括機械、組合或高階程序、邏輯或物件導向程式設計語言),以與電腦系統溝通。在任何情形中,該語言係為編譯或解釋語言。

例如,當由電腦軟體指令序列來實施時,本發明實施例的各類功能與步驟可藉由在適當數位訊號處理硬體中執行的多線軟體指令序列來實施,在該情形中,該等實施例的各類裝置、步驟與功能係對應該等軟體指令部分。

各此電腦程式較佳地被儲存在或下載到可由一般或特殊目的可程式電腦來讀取的儲存媒體或裝置(例如,固態記憶體或媒體,或者磁性或光學媒體),以用於當該儲存媒體或裝置藉由該電腦系統所讀取以施行在本文中所說明之程序時來組態與操作該電腦。本發明系統亦以電腦可讀取儲存媒體來實施、以電腦程式來組態(亦即,儲存),在此如此組態的儲存媒體會導致電腦系統以具體且預定的方式操作,以施行在本文中所說明的功能。

本發明的許多實施例已經被說明。然而,將理解的是,各類修改可在不背離本發明精神與範圍之下進行。根據以上學說,本發明的許多修改與變化是可能的。要理解的是,在附加申請專利範圍的範圍內,本發明可能以除了

在本文中所具體說明以外的方式來實施。

【符號說明】

- 100：編碼器
- 101：解碼器
- 102：音頻狀態驗證器
- 103：響度處理台
- 104：音頻流選擇台
- 105：編碼器
- 106：元資料產生台
- 107：填充器/格式器台
- 108：對話響度測量子系統
- 109：訊框緩衝器
- 110：訊框緩衝器
- 111：剖析器
- 120：驗證器
- 150：已編碼音頻傳送子系統
- 152：解碼器
- 200：解碼器
- 201：訊框緩衝器
- 202：音頻解碼器
- 203：音頻狀態驗證台（驗證器）
- 204：控制位元產生台
- 205：剖析器

300 : 後處理器

301 : 訊框緩衝器

申請專利範圍

1. 一種音頻處理裝置，用於解碼已編碼音頻位元流，該音頻處理裝置包含：

輸入緩衝器，其儲存該已編碼音頻位元流之至少一部分，該已編碼音頻位元流包括音頻資料與元資料容器；

位元流剖析器，其剖析該音頻資料；以及

解碼器，其解碼該音頻資料，

其中該已編碼音頻位元流被切成一或多個訊框，每個訊框包括：

同步資訊段，包括訊框同步字，

位元流資訊段，接在該同步資訊段之後，該位元流資訊包括音頻元資料，

可選的額外位元流資訊段，位在該位元流資訊段的末端，

多達六個方塊的音頻資料，接在該位元流資訊段之後，

輔助資訊段，接在該多達六個方塊的音頻資料之後，

錯誤校正字，接在該輔助資訊段之後，以及

一或多個可選的跨越(skip)欄位，包含殘留在該訊框中的任何未使用的空間，

其中該一或多個訊框的其中至少一個包括該元資料容器，該元資料容器位在選自由該一或多個可選的跨越欄位、該額外位元流資訊段、該輔助資訊段或其組合所組成

之群組的保留資料空間中，

其中該元資料容器包括：

標頭，識別該元資料容器之開頭，該標頭包括同步字，其後接著指定該元資料容器之長度的長度欄位，

語法版本欄位，接在該標頭之後，該語法版本欄位指定該元資料容器的語法版本，

一或多個元資料酬載，接在該版本欄位之後，每個元資料酬載包括唯一地識別該元資料酬載的識別符，該元資料酬載之後接著該元資料酬載的元資料，以及

保護資料，接在該一或多個元資料酬載之後，該保護資料可被用來鑑定或驗證該元資料容器或者在該元資料容器之內的該一或多個元資料酬載。

2. 如申請專利範圍第 1 項之音頻處理裝置，其中該同步字為 16 位元欄位，具有 0x5838 的值。

3. 如申請專利範圍第 1 項之音頻處理裝置，其中該一或多個元資料酬載包括節目響度酬載，且該響度酬載包括對話頻道欄位，該對話頻道欄位由指示該音頻資料的左、右、或中央頻道是否包含口語對話的 3 位元欄位組成。

4. 如申請專利範圍第 1 項之音頻處理裝置，其中該一或多個元資料酬載包括節目響度酬載，且該響度酬載包括響度規則類型欄位，該響度規則類型欄位由指示哪個響度規則標準被用來計算與該音頻資料相關聯的節目響度的 4 位元欄位組成。

5. 如申請專利範圍第 1 項之音頻處理裝置，其中該一或多個元資料酬載包括節目響度酬載，其中該節目響度酬載包括響度校正類型，該響度校正類型由指示該音頻資料是否已用無限預見或以檔案為基礎的響度校正處理來校正的 1 位元欄位組成。

6. 如申請專利範圍第 1 項之音頻處理裝置，其中該已編碼音頻位元流為 AC-3 位元流或 E-AC-3 位元流。

7. 一種用於解碼已編碼音頻位元流的方法，該方法包含以下步驟：

接收該已編碼音頻位元流之至少一部分，該已編碼音頻位元流包括音頻資料與元資料容器；

剖析該音頻資料；以及

解碼該音頻資料，

其中該已編碼音頻位元流被切成一或多個訊框，每個訊框包括：

同步資訊段，包括訊框同步字，

位元流資訊段，接在該同步資訊段之後，該位元流資訊包括音頻元資料，

可選的額外位元流資訊段，位在該位元流資訊段的末端，

多達六個方塊的音頻資料，接在該位元流資訊段之後，

輔助資訊段，接在該多達六個方塊的音頻資料之後，

錯誤校正字，接在該輔助資訊段之後，以及
一或多個可選的跨越欄位，包含殘留在該訊框中的任何未使用的空間，

其中該一或多個訊框的其中至少一個包括該元資料容器，該元資料容器位在選自由該一或多個可選的跨越欄位、該額外位元流資訊段、該輔助資訊段或其組合所組成之群組的保留資料空間中，

其中該元資料容器包括：

標頭，識別該元資料容器之開頭，該標頭包括同步字，其後接著指定該元資料容器之長度的長度欄位，

語法版本欄位，接在該標頭之後，該語法版本欄位指定該元資料容器的語法版本，

一或多個元資料酬載，接在該版本欄位之後，每個元資料酬載包括唯一地識別該元資料酬載的識別符，該元資料酬載之後接著該元資料酬載的元資料，以及

保護資料，接在該一或多個元資料酬載之後，該保護資料可被用來鑑定或驗證該元資料容器或者在該元資料容器之內的該一或多個元資料酬載。



圖 1

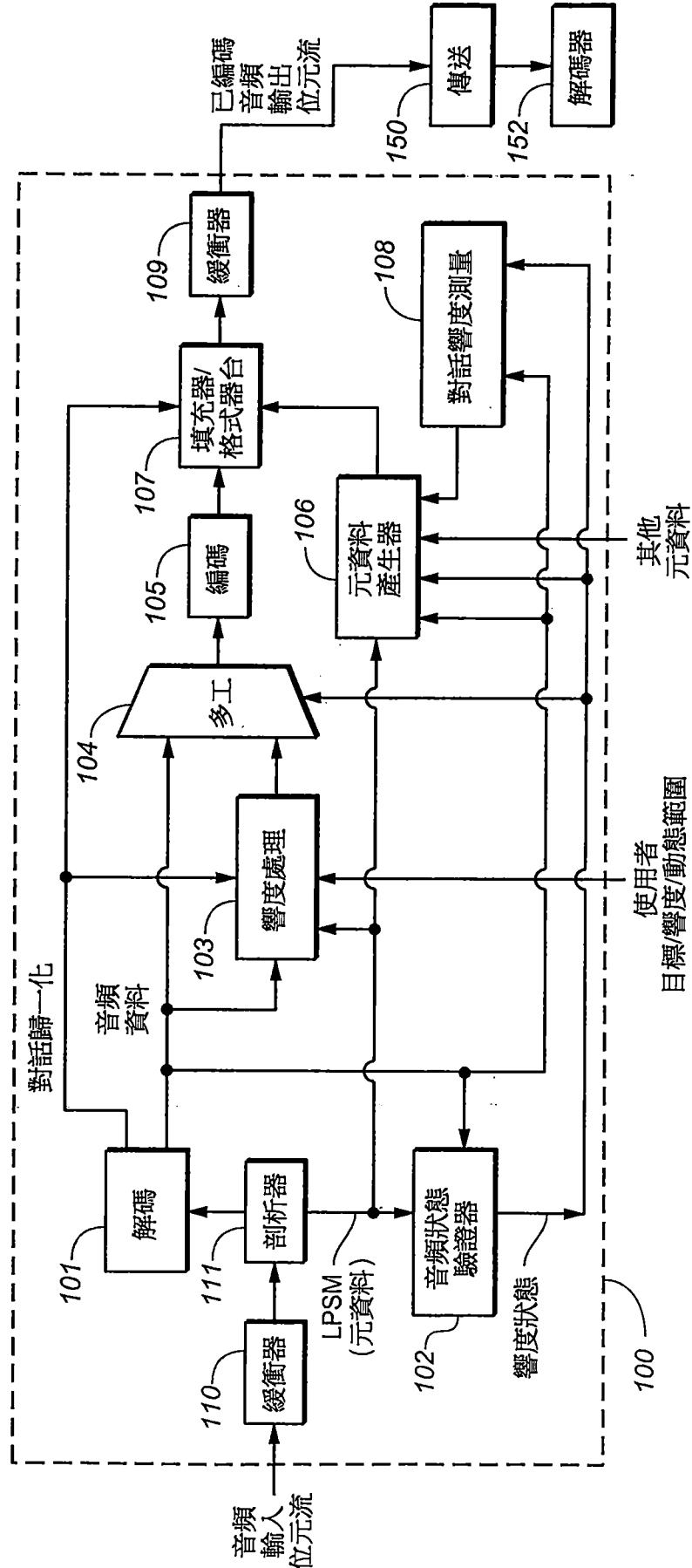


圖 2

圖 2

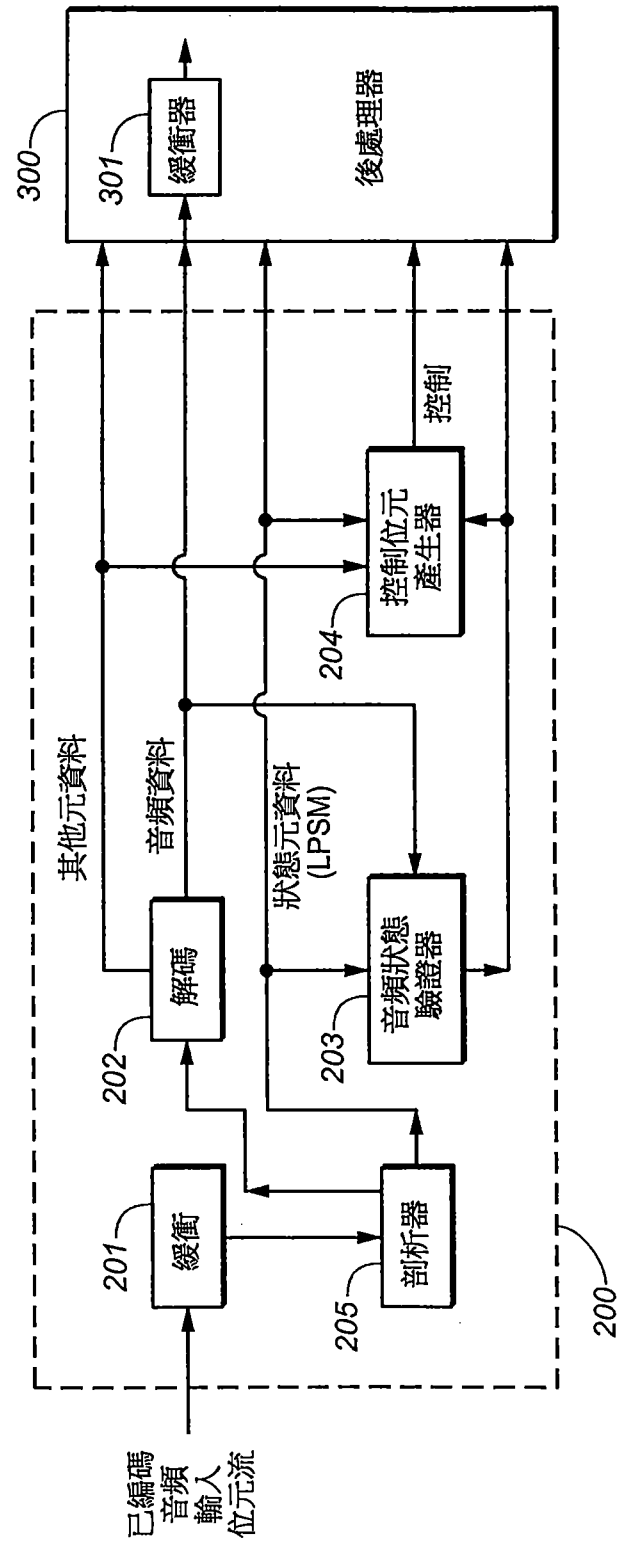


圖 3

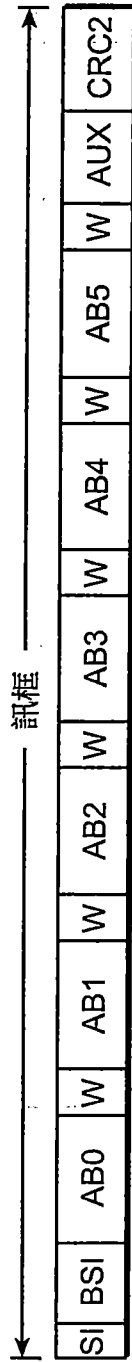


圖 4

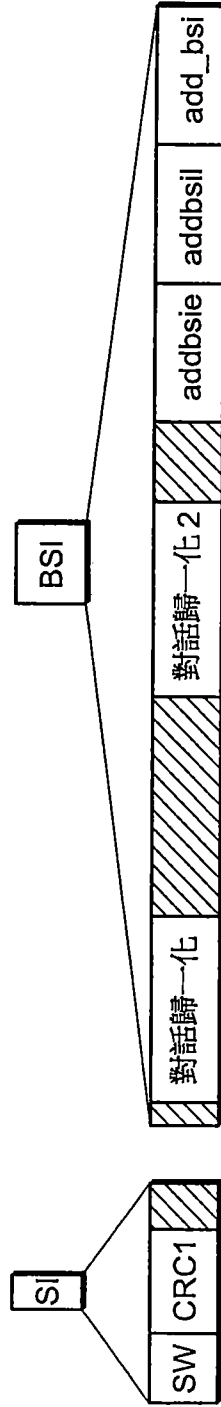


圖 5

圖 6

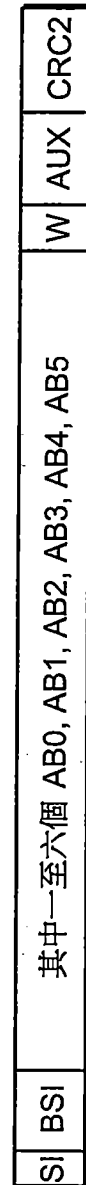


圖 7

訊框 #	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	
到邊界的訊框	-16	-15	-14	-13	-12	-11	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	B
節目邊界旗標內容	100011	0	0	0	0	0	0	0	10011	0	0	0	1011	0	111	0	B
存在	1	0	0	0	0	0	0	0	1	0	0	0	1	0	1	0	B
訊框計數器	0001->16								001->8				01->4		1->2		B
計數 (+) / 倒數 (-)	1								1						1		B

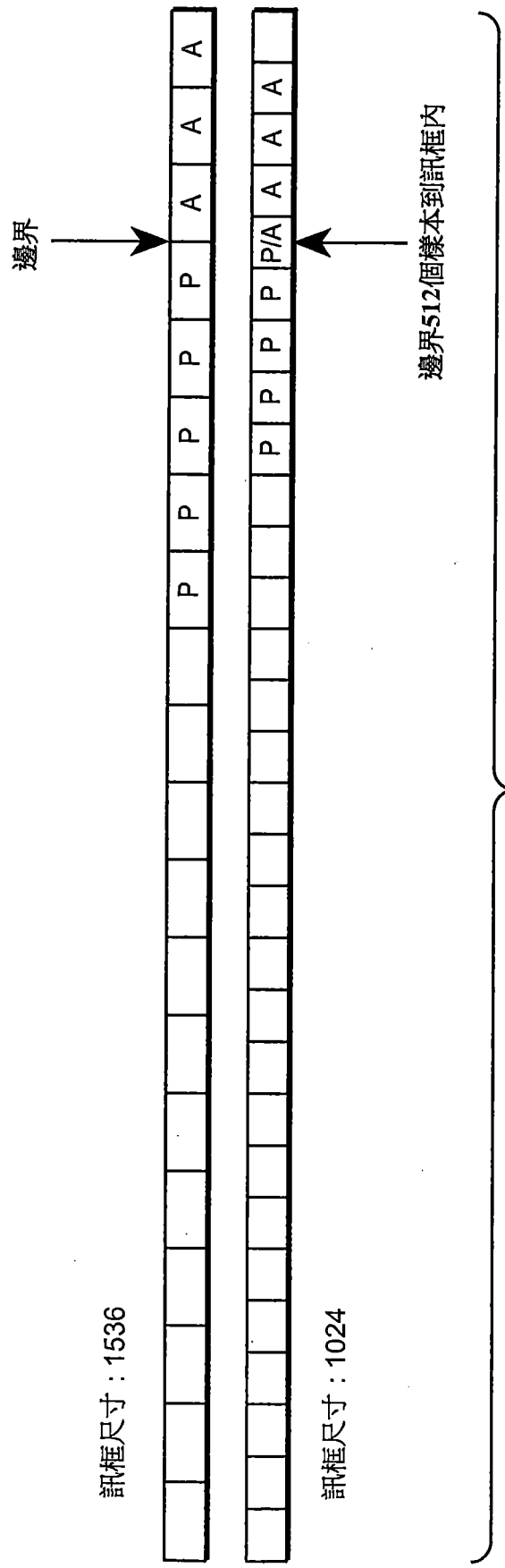
(倒數至節目邊界之節目邊界旗標的語法)

圖 8

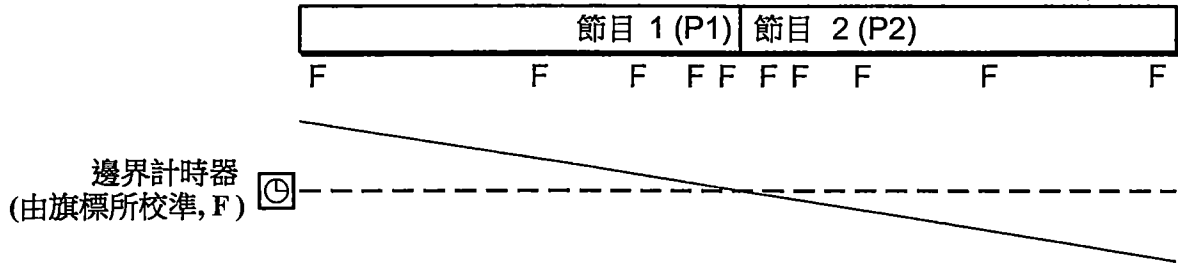
訊框 #	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
到邊界的訊框	B	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
節目邊界旗標內容	B	0	110	0	1010	0	0	10010	0	0	0	0	0	0	0	10010
存在	B	0	1	0	0	0	0	1	0	0	0	0	0	0	0	
訊框計數器	B	1->2		01->4				001->8								0001->16
計數 (+) / 倒數 (-)	B	0		0				0								0

(自節目開始計數之節目邊界旗標的語法)

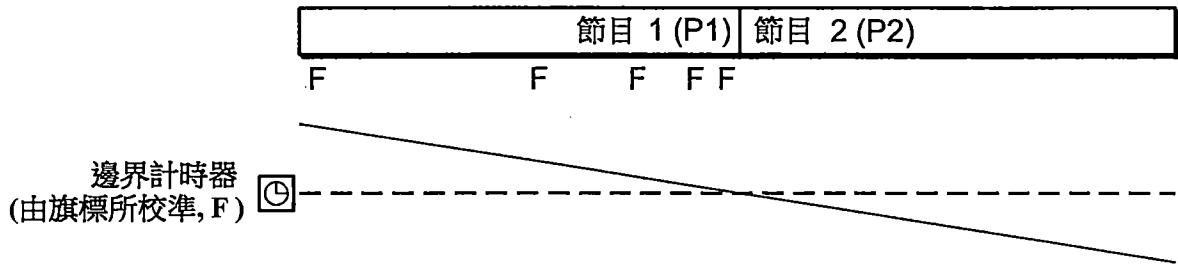
圖 9



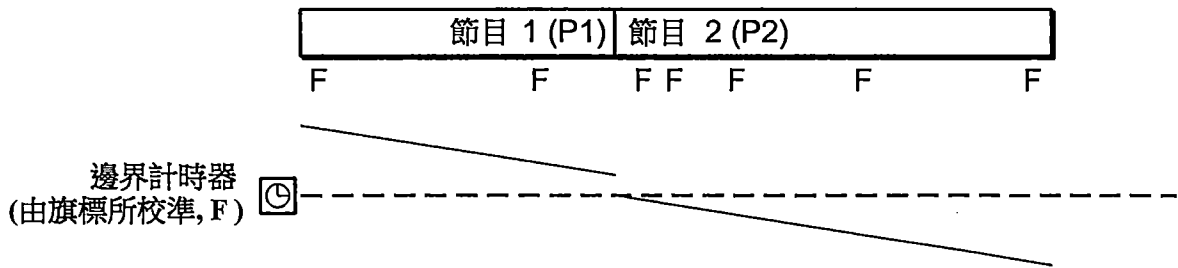
情境 1：P1與P2兩者皆具有旗標



情境 2：只有P1具有旗標



情境 3：被剪接之P1結束



情境 4：P1結束、被剪接之P2開頭

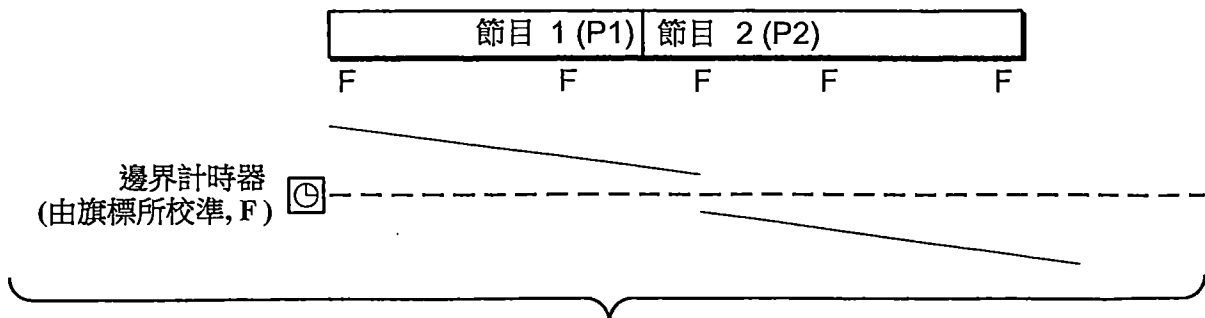


圖 11

發明摘要

※申請案號：

※申請日：

※IPC 分類：

【發明名稱】(中文/英文)

使用位在保留資料空間中的元資料容器解碼已編碼音
頻位元流

Decoding of encoded audio bitstream with metadata container located in
reserved data space

【中文】

本發明揭露用來產生一已編碼音頻位元流的設備與方法，其包括藉由包括節目響度元資料與音頻資料於該位元流中，以及可選地亦包括節目邊界元資料於該位元流的至少一個片段（例如，訊框）中。其他態樣係為用來解碼此類位元流的設備與方法，例如包括藉由施行由該位元流所指示出之一音頻節目的該音頻資料的自適應響度處理，或者此類音頻節目之元資料及/或音頻資料的鑑定及/或驗證。另一態樣係為一音頻處理單元（例如，一編碼器、解碼器或後處理器），其被組態（例如，程式化）以施行該方法之任何實施例或者其包括儲存根據該方法之任何實施例所產生之一音頻位元流之至少一個訊框的一緩衝記憶體。

【 英文 】

Apparatus and methods for generating an encoded audio bitstream, including by including program loudness metadata and audio data in the bitstream, and optionally also program boundary metadata in at least one segment (e.g., frame) of the bitstream. Other aspects are apparatus and methods for decoding such a bitstream, e.g., including by performing adaptive loudness processing of the audio data of an audio program indicated by the bitstream, or authentication and/or validation of metadata and/or audio data of such an audio program. Another aspect is an audio processing unit (e.g., an encoder, decoder, or post-processor) configured (e.g., programmed) to perform any embodiment of the method or which includes a buffer memory which stores at least one frame of an audio bitstream generated in accordance with any embodiment of the method.

發明專利說明書

(本說明書格式、順序，請勿任意更動)

【發明名稱】(中文/英文)

使用位在保留資料空間中的元資料容器解碼已編碼音
頻位元流

Decoding of encoded audio bitstream with metadata container located in
reserved data space

【技術領域】

本發明關於音頻訊號處理，且更特別地關於音頻資料位元流之編碼與解碼，其具有指示出音頻內容之響度處理狀態以及由該等位元流所指示之音頻節目邊界的位置的元資料。本發明的一些實施例將呈已知為 AC-3、增強 AC-3、或 E-AC-3 或杜比 E 之該等格式之其中一個的音頻資料產生或解碼。

【先前技術】

杜比、杜比數位、杜比數位 Plus、與杜比 E 係為杜比實驗室特許公司的商標。杜比實驗室提供個別已知為杜比數位與杜比數位 Plus 之 AC-3 與 E-AC-3 的專屬實施程序。

音頻資料處理單元基本上以盲目的方式操作，並且沒有注意到在接收資料以前所發生之音頻資料的處理歷史。這或許可在單一個體進行全部的音頻資料處理與編碼，以

用於各種目標媒體顯像裝置，同時目標媒體顯像裝置進行已編碼音頻資料的全部解碼與顯像的處理架構中運行。不過，此盲目處理在複數個音頻處理單元被分散在不同網路或者串聯地放置（亦即，鏈）以及被期待最佳施行它們各別類型的音頻處理的情況下沒有（或者一點都沒有）運行良好。例如，一些音頻資料可被編碼以用於高性能媒體系統，並且必須被轉換成適合用於沿著一媒體處理鏈之行動裝置的一縮減形式。據此，音頻處理單元可非必要地施行一種處理於已經被施行的音頻資料上。例如，音量分級單元可施行處理於輸入音頻段落（clip）上，不管相同或相似的音量分級是否已經被事先施行於輸入音頻段落上。因此，音量分級單元甚至在不必要時仍施行分級。此非必要處理亦會在顯現音頻資料的內容的同時，導致具體特徵的退化與/或移除。

典型的音頻資料流包括指示出該音頻內容之至少一個特徵的音頻內容（例如，音頻內容的一或多個頻道）與元資料兩者。例如，在 AC-3 位元流中，有數個音頻元資料參數，其係特別打算使用來改變被傳送到聆聽環境之節目的聲音。該等元資料參數的其中一個為對話歸一化（DIALNORM）參數，其係打算指出發生音頻節目之對話的平均位準，並且被使用來決定音頻播放訊號位準。

在包含一序列不同音頻節目片段之位元流的播放期間內（各個均具有不同的 DIALNORM 參數），AC-3 解碼器使用各片段的 DIALNORM 參數來施行一種響度處理，其